(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 August 2006 (10.08.2006)

PCT

(10) International Publication Number
WO 2006/083800 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2006/003298

(22) International Filing Date: 31 January 2006 (31.01.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/048,627 31 January 2005 (31.01.2005) US

(71) Applicant *(for all designated States except US)*: **UNIVERSITY OF IOWA RESEARCH FOUNDATION** [US/US]; 100 Oakdale Campus, 214 Technology Innovation Center, Iowa City, IA 52242 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **DAVIDSON, Beverly, L.** [US/US]; 3640 Johnson Way N.E., North Liberty, IA 52242 (US). **HARPER, Scott** [US/US]; 1 Warwick Circle, Iowa City, IA 52240-2847 (US).

(74) Agents: **VIKSNINS, Ann, S.** et al.; Viksnins Harris & Padys PLLP, P.O. Box 111098, St. Paul, MN 55111-1098 (US).

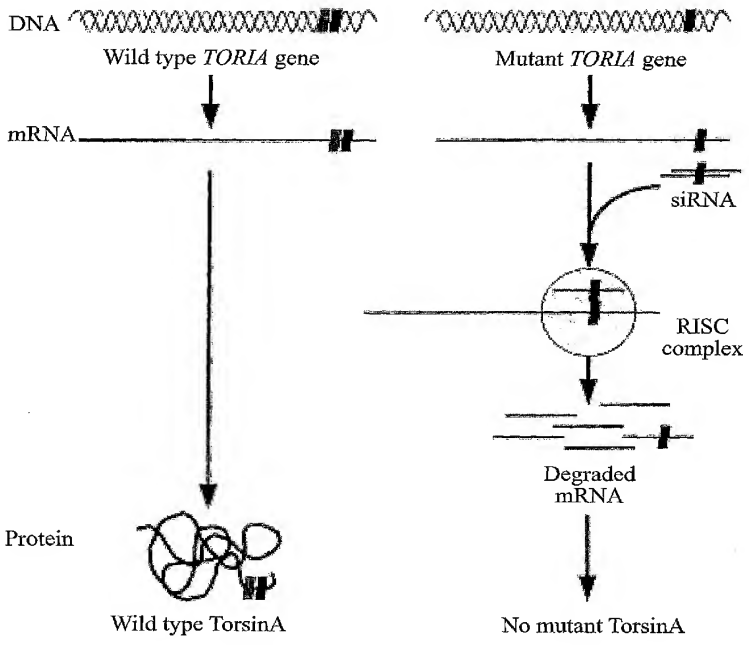
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** NUCLEIC ACID SILENCING OF HUNTINGTON'S DISEASE GENE



(57) **Abstract:** The present invention is directed to small interfering RNA molecules (siRNA) targeted against a Huntington's Disease gene, and methods of using these siRNA molecules.

WO 2006/083800 A2

NUCLEIC ACID SILENCING OF HUNTINGTON'S DISEASE GENE

Claim of Priority

This patent application claims priority to U.S. Patent Application No. 11/048,627 filed on January 31, 2005 which is a continuation-in-part application of U.S. Application Serial No. 10/859,751 filed on June 2, 2004, which is a continuation-in-part of International PCT Application No. PCT/US03/16887 filed on May 26, 2003, which is a continuation-in-part of application U.S. Application Serial No. 10/430,351 filed on May 5, 2003, which is a continuation of U.S. Application Serial No. 10/322,086 filed on December 17, 2002, which is a continuation-in-part application of U.S. Application Serial No. 10/212,322, filed August 5, 2002. The instant application claims the benefit of all the listed applications, which are hereby incorporated by reference herein in their entireties, including the drawings.

Statement Regarding Federally Sponsored Research Or Development

Work relating to this application was supported by a grant from the National Institutes of Health, NS044494. The government may have certain rights in the invention.

Background of the Invention

Double-stranded RNA (dsRNA) can induce sequence-specific posttranscriptional gene silencing in many organisms by a process known as RNA interference (RNAi). However, in mammalian cells, dsRNA that is 30 base pairs or longer can induce sequence-nonspecific responses that trigger a shut-down of protein synthesis. Recent work suggests that RNA fragments are the sequence-specific mediators of RNAi (Zamore *et al.*, 2000, Cell, 101, 25-33; Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188). Interference of gene expression by these small interfering RNA (siRNA) is now recognized as a naturally occurring strategy for silencing genes in *C. elegans*, *Drosophila*, plants, and in mouse

embryonic stem cells, oocytes and early embryos (Cogoni *et al.*, 1994; Baulcombe, 1996; Kennerdell, 1998; Timmons, 1998; Waterhouse *et al.*, 1998; Wianny and Zernicka-Goetz, 2000; Yang *et al.*, 2001; Svoboda *et al.*, 2000). In mammalian cell culture, an siRNA-mediated reduction in gene expression has been accomplished only by transfecting cells with synthetic RNA oligonucleotides (Caplan *et al.*, 2001; Elbashir *et al.*, 2001).

Summary of the Invention

This invention relates to compounds, compositions, and methods useful for modulating Huntington's Disease (also referred to as huntingtin, htt, or HD) gene expression using short interfering nucleic acid (siNA) molecules. This invention also relates to compounds, compositions, and methods useful for modulating the expression and activity of other genes involved in pathways of HD gene expression and/or activity by RNA interference (RNAi) using small nucleic acid molecules. In particular, the instant invention features small nucleic acid molecules, such as short interfering nucleic acid (siNA), short interfering RNA (siRNA), double-stranded RNA (dsRNA), micro-RNA (miRNA), and short hairpin RNA (shRNA) molecules and methods used to modulate the expression HD genes. A siNA of the instant invention can be chemically synthesized, expressed from a vector or enzymatically synthesized.

In one embodiment, the present invention provides an AAV-1 expressed siRNA comprising an isolated first strand of RNA of 15 to 30 nucleotides in length and an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first or second strand comprises a sequence that is complementary to a nucleotide sequence encoding a mutant Huntington's Disease protein, wherein at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex under physiological conditions, and wherein the siRNA silences the expression of the nucleotide sequence encoding the mutant Huntington's Disease protein in the cell. In one embodiment, the first or second strand comprises a sequence that is complementary to both a mutant and wild-type Huntington's disease allele, and

the siRNA silences the expression of the nucleotide sequence encoding the mutant Huntington's Disease protein and wild-type Huntington's Disease protein in the cell.

In one embodiment, the present invention provides an AAV-1 expressed siRNA comprising an isolated first strand of RNA of 15 to 30 nucleotides in length and an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first or second strand comprises a sequence that is complementary to both a nucleotide sequence encoding a wild-type and mutant Huntington's Disease protein, wherein at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex under physiological conditions, and wherein the siRNA silences the expression of the nucleotide sequence encoding the wild-type and mutant Huntington's Disease protein in the cell. In one embodiment, an AAV-1 vector of the invention is a psuedotyped rAAV-1 vector.

In one embodiment, the present invention provides a mammalian cell containing an isolated first strand of RNA for example corresponding to SEQ ID NO:60, SEQ ID NO:63, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, or SEQ ID NO:88, and an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first strand contains a sequence that is complementary to a nucleotide sequence encoding a Huntington's Disease protein (htt), such as wherein at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex for example under physiological conditions, and wherein the siRNA silences the expression of the Huntington's Disease (HD) gene in the cell, for example by targeting the cleavage of RNA encoded by the HD gene or via translational blocking of the HD gene expression. SEQ ID NO:60 through SEQ ID NO:89 are all represented herein as DNA sequences. However, as used herein when a claim indicates an RNA "corresponding to" it is meant the RNA that has the same sequence as the DNA, except that uracil is substituted for thymine. For example, SEQ ID NO:61 is 5'-

GAAGCTTG-3', and the RNA corresponding to this sequence is 5'-GAAGCUUG-3' (SEQ ID NO: 58).

The present invention provides a method of suppressing the accumulation of huntingtin in a cell by introducing a ribonucleic acid (RNA) described above into the cell in an amount sufficient to suppress accumulation of huntingtin in the cell. In certain embodiments, the accumulation of huntingtin is suppressed by at least 10%. The accumulation of huntingtin is suppressed by at least 10%, 20%, 30%, 40%, 50%, 60%, 70% 80%, 90% 95%, or 99%.

The present invention provides a method to inhibit expression of a huntingtin gene in a cell by introducing a ribonucleic acid (RNA) described above into the cell in an amount sufficient to inhibit expression of the huntingtin, and wherein the RNA inhibits expression of the huntingtin gene. The huntingtin is inhibited by at least 10%, 20%, 30%, 40%, 50%, 60%, 70% 80%, 90% 95%, or 99%.

The present invention provides a method to inhibit expression of a huntingtin gene in a mammal (*e.g.*, a human) by (a) providing a mammal containing a neuronal cell, wherein the neuronal cell contains the huntingtin gene and the neuronal cell is susceptible to RNA interference, and the huntingtin gene is expressed in the neuronal cell; and (b) contacting the mammal with a ribonucleic acid (RNA) or a vector described above, thereby inhibiting expression of the huntingtin gene. In certain embodiments, the accumulation of huntingtin is suppressed by at least 10%. The huntingtin is inhibited by at least 10%, 20%, 30%, 40%, 50%, 60%, 70% 80%, 90% 95%, or 99%. In certain embodiments, the cell located in vivo in a mammal.

The present invention also provides a method to inhibit expression of a protein associated with the neurodegenerative disease, such as huntingtin, in a mammal in need thereof, by introducing the vector encoding a miRNA described above into a cell in an amount sufficient to inhibit expression of the protein associated with the neurodegenerative disease, wherein the RNA inhibits expression of the protein associated with the neurodegenerative disease. The

protein is inhibited by at least 10%, 20%, 30%, 40%, 50%, 60%, 70% 80%, 90% 95%, or 99%.

The present invention provides a method to inhibit expression of huntingtin in a mammal in need thereof by (a) providing a mammal containing a neuronal cell, wherein the neuronal cell contains the huntingtin gene and the neuronal cell is susceptible to RNA interference, and the huntingtin gene is expressed in the neuronal cell; and (b) contacting the mammal the vector encoding a miRNA described above, thereby inhibiting expression of the huntingtin gene. The huntingtin is inhibited by at least 10%, 20%, 30%, 40%, 50%, 60%, 70% 80%, 90% 95%, or 99%.

In one embodiment, the invention features siRNA duplexes where the first and/or second strand of the duplex further include a 3' overhang region, a 5' overhang region, or both 3' and 5' overhang regions, and the overhang region (or regions) can be from 1 to 10 nucleotides in length. As used herein, the term "overhang region" means a portion of the RNA that does not bind with the second strand. Further, the first strand and the second strand encoding the siRNA duplex can be operably linked by means of an RNA loop strand to form a hairpin structure comprising a duplex structure and a loop structure. Such siRNAs with hairpin stem-loop structure are referred to sometimes as short hairpin RNAs or shRNAs. This loop structure, if present may be from 4 to 10 nucleotides or longer in length. In one embodiment, the loop structure corresponds to SEQ ID NO:58. In one embodiment, the first strand corresponds to SEQ ID NO:56 and the second strand corresponds to SEQ ID NO:57.

The reference to siRNAs herein is meant to include shRNAs and other small RNAs that can or are capable of modulating the expression of HD gene, for example via RNA interference. Such small RNAs include without limitation, shRNAs and miroRNAs (miRNAs).

The present invention also provides a mammalian cell containing an expression cassette encoding an isolated first strand of RNA corresponding to, for example, SEQ ID NO:56 or SEQ ID NO:57, and encoding an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first or second

strand comprises a sequence that is complementary to a nucleotide sequence encoding a Huntington's Disease protein (htt), for example wherein at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex for example under physiological conditions, and wherein the siRNA silences the expression of the Huntington's Disease gene in the cell, for instance by targeting the cleavage of RNA encoded by the HD gene or via translational blocking of the HD gene expression. The expression cassette may further include a promoter, such as a regulatable promoter or a constitutive promoter. Examples of suitable promoters include without limitation a pol II promoter such as cytomegalovirus (CMV), Rous Sarcoma Virus (RSV), pol III promoters such as U6, and any other pol II or pol III promoter as is known in the art. The expression cassette may further optionally include a marker gene, such as a stuffer fragment comprising a marker gene. The expression cassette may be contained in a vector, such as an adenoviral, lentiviral, adeno-associated viral (AAV), poliovirus, HSV, or murine Maloney-based viral vector. In one embodiment, the first strand corresponds to SEQ ID NO:56 and the second strand corresponds to SEQ ID NO:57.

The present invention provides a small interfering RNA (siRNA) containing a first strand of RNA corresponding to for example SEQ ID NO:56 or SEQ ID NO:57, and a second strand of RNA of 15 to 30 nucleotides in length, wherein the first or second strand comprises a sequence that is complementary to a nucleotide sequence encoding a Huntington's Disease protein (htt), for example wherein at least 12 nucleotides of the first and second strands are complementary to each other and form an siRNA duplex under physiological conditions, wherein the duplex is between 15 and 30 base pairs in length, and wherein the siRNA silences the expression of the Huntington's Disease gene in the cell, for instance via RNA interference.

The present invention provides a method of performing Huntington's Disease gene silencing in a mammal by administering to the mammal an expression cassette encoding an isolated first strand of RNA corresponding to for example SEQ ID NO:56 or SEQ ID NO:57, and encoding an isolated second

strand of RNA of 15 to 30 nucleotides in length, wherein the first or second strand comprises a sequence that is complementary to a nucleotide sequence encoding a Huntington's Disease protein (htt), for example wherein at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex under physiological conditions, and wherein the expression of the siRNA from the expression cassette silences the expression of the Huntington's Disease gene in the mammal, for instance via RNA interference.

The present invention provides an isolated RNA comprising for example SEQ ID NO:59 that functions in RNA interference to a sequence encoding a mutant Huntington's Disease protein (htt).

The present invention provides an isolated RNA duplex comprising a first strand of RNA corresponding to for example SEQ ID NO:56 and a second strand of RNA corresponding to for example by SEQ ID NO:57. The first and/or second strand optionally further include a 3' overhang region, a 5' overhang region, or both 3' and 5' overhang regions, and the overhang region (or regions) can be from 1 to 10 nucleotides in length. Further, the first strand and the second strand can be operably linked by means of an RNA loop strand to form a hairpin structure comprising a duplex structure and a loop structure. This loop structure, if present may be from 4 to 10 nucleotides. In one embodiment, the loop structure corresponds to SEQ ID NO:58 or a portion thereof.

The present invention provides a vector, such as an AAV vector, comprising two expression cassettes, a first expression cassette comprising a nucleic acid encoding the first strand of the RNA duplex corresponding to for example SEQ ID NO:56 and a second expression cassette comprising a nucleic acid encoding the second strand of the RNA duplex corresponding to for example SEQ ID NO:57. The present invention also provides a cell containing this vector. In one embodiment, the cell is a mammalian cell.

The present invention provides a mammalian cell containing an isolated first strand of RNA of 15 to 30 nucleotides in length, and an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first strand contains

a sequence that is complementary to for example at least 15 nucleotides of RNA encoded by a targeted gene of interest (for example the HD gene), wherein for example at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex for example under physiological conditions, and wherein the siRNA silences (for example via RNA interference) only one allele of the targeted gene (for example the mutant allele of HD gene) in the cell. The duplex of the siRNA may be between 15 and 30 base pairs in length. The two strands of RNA in the siRNA may be completely complementary, or one or the other of the strands may have an "overhang region" or a "bulge region" (*i.e.*, a portion of the RNA that does not bind with the second strand or where a portion of the RNA sequence is not complementary to the sequence of the other strand). These overhangs may be at the 3' end or at the 5' region, or at both 3' and 5' ends. Such overhang regions may be from 1 to 10 (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10) or more nucleotides in length. The bulge regions may be at the ends or in the internal regions of the siRNA duplex. Such bulge regions may be from 1-5 (*e.g.*, 1, 2, 3, 4, 5) or more nucleotides long. Such bulge regions may be the bulge regions characteristics of miRNAs. In the present invention, the first and second strand of RNA may be operably linked together by means of an RNA loop strand to form a hairpin structure to form a "duplex structure" and a "loop structure." These loop structures may be from 4 to 10 (*e.g.*, 4, 5, 6, 7, 8, 9, 10) or more nucleotides in length. For example, the loop structure may be 4, 5 or 6 nucleotides long.

The present invention also provides a mammalian cell that contains an expression cassette encoding an isolated first strand of RNA of 15 to 30 nucleotides in length, and an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first strand contains a sequence that is complementary to for example at least 15 contiguous nucleotides of RNA encoded by a targeted gene of interest (for example the HD gene), wherein for example at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex, for example under physiological conditions, and wherein the siRNA silences (for

example via RNA interference) only one allele of the targeted gene (for example the mutant allele of HD gene) in the cell. These expression cassettes may further contain a promoter. Such promoters can be regulatable promoters or constitutive promoters. Examples of suitable promoters include a CMV, RSV, pol II or pol III promoter. The expression cassette may further contain a polyadenylation signal, such as a synthetic minimal polyadenylation signal. The expression cassette may further contain a marker gene. The expression cassette may be contained in a vector. Examples of appropriate vectors include adenoviral, lentiviral, adeno-associated viral (AAV), poliovirus, HSV, or murine Maloney-based viral vectors. In one embodiment, the vector is an adenoviral vector or an adeno-associated viral vector.

In the present invention, the alleles of the targeted gene may differ by seven or fewer nucleotides (*e.g.*, 7, 6, 5, 4, 3, 2 or 1 nucleotides). For example the alleles may differ by only one nucleotide. Examples of targeted gene transcripts include transcripts encoding a beta-glucuronidase, TorsinA, Ataxin-3, Tau, or huntingtin. The targeted genes and gene products (*i.e.*, a transcript or protein) may be from different species of organisms, such as a mouse allele or a human allele of a target gene.

The present invention also provides an isolated RNA duplex containing a first strand of RNA and a second strand of RNA, wherein the first strand contains for example at least 15 nucleotides complementary to mutant *TorsinA* represented for example by SEQ ID NO:55, and wherein the second strand is complementary to for example at least 12 contiguous nucleotides of the first strand. In one embodiment of the invention (mutA-si), the first strand of RNA corresponds to for example SEQ ID NO:49 and the second strand of RNA corresponds to for example SEQ ID NO:50. In an alternative embodiment (mutB-si), the first strand of RNA corresponds to for example SEQ ID NO:51 and the second strand of RNA corresponds to for example SEQ ID NO:52. In another embodiment (mutC-si), the first strand of RNA corresponds to for example SEQ ID NO:53 and second strand of RNA corresponds to for example SEQ ID NO:54. As used herein the term "encoded by" means that the DNA

sequence is transcribed into the RNA of interest. This term is used in a broad sense, similar to the term "comprising" in patent terminology. For example, the statement "the first strand of RNA is encoded by SEQ ID NO:49" means that the first strand of RNA sequence corresponds to the DNA sequence indicated in SEQ ID NO:49, but may also contain additional nucleotides at either the 3' end or at the 5' end of the RNA molecule.

The present invention further provides an RNA duplex containing a first strand of RNA and a second strand of RNA, wherein the first strand contains for example at least 15 contiguous nucleotides complementary to mutant *Ataxin-3* transcript encoded by SEQ ID NO:8, and wherein the second strand is complementary to for example at least 12 contiguous nucleotides of the first strand. In one embodiment (siC7/8), the first strand of RNA is encoded by SEQ ID NO:19 and the second strand of RNA is encoded by SEQ ID NO: 20. In another embodiment (siC10), the first strand of RNA is encoded by SEQ ID NO:21 and the second strand of RNA is encoded by SEQ ID NO:22.

The present invention further provides an RNA duplex containing a first strand of RNA and a second strand of RNA, wherein the first strand contains for example at least 15 contiguous nucleotides complementary to mutant Tau transcript for example encoded by SEQ ID NO:39 (siA9/C12), and wherein the second strand is complementary to at least 12 contiguous nucleotides of the first strand. The second strand may be encoded for example by SEQ ID NO:40.

The RNA duplexes of the present invention are between 15 and 30 base pairs in length. For example they may be between 19 and 25 base pairs in length or 19-27 base-pairs in length. As discussed above the first and/or second strand further may optionally comprise an overhang region. These overhangs may be at the 3' end or at the 5' overhang region, or at both 3' and 5' ends. Such overhang regions may be from 1 to 10 nucleotides in length. The RNA duplex of the present invention may optionally include nucleotide bulge regions. The bulge regions may be at the ends or in the internal regions of the siRNA duplex. Such bulge regions may be from 1-5 nucleotides long. Such bulge regions may be the bulge regions characteristics of miRNAs. In the present invention, the first and

second strand of RNA may be operably linked together by means of an RNA loop strand to form a hairpin structure to form a "duplex structure" and a "loop structure." These loop structures may be from 4 to 10 nucleotides in length. For example, the loop structure may be 4, 5 or 6 nucleotides long.

In the present invention, an expression cassette may contain a nucleic acid encoding at least one strand of the RNA duplex described above. Such an expression cassette may further contain a promoter. The expression cassette may be contained in a vector. These cassettes and vectors may be contained in a cell, such as a mammalian cell. A non-human mammal may contain the cassette or vector. The vector may contain two expression cassettes, the first expression cassette containing a nucleic acid encoding the first strand of the RNA duplex, and a second expression cassette containing a nucleic acid encoding the second strand of the RNA duplex.

In one embodiment, the present invention further provides a method of performing gene silencing in a mammal or mammalian cell by administering to the mammal an isolated first strand of RNA of about 15 to about 30 nucleotides (for example 19-27 nucleotides) in length, and an isolated second strand of RNA of 15 to 30 nucleotides (for example 19-27 nucleotides) in length, wherein the first strand contains for example at least 15 contiguous nucleotides complementary to a targeted gene of interest (such as HD gene), wherein for example at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex for example under physiological conditions, and wherein the siRNA silences only one or both alleles of the targeted gene (for example the wild type and mutant alleles of HD gene) in the mammal or mammalian cell. In one example, the gene is a beta-glucuronidase gene. The alleles may be murine-specific and human-specific alleles of beta-glucuronidase. Examples of gene transcripts include an RNA transcript complementary to *TorsinA*, *Ataxin-3*, *huntingtin* or *Tau*. The targeted gene may be a gene associated with a condition amenable to siRNA therapy. For example, the condition amenable to siRNA therapy could be a disabling neurological disorder.

"Neurological disease" and "neurological disorder" refer to both hereditary and sporadic conditions that are characterized by nervous system dysfunction, and which may be associated with atrophy of the affected central or peripheral nervous system structures, or loss of function without atrophy. A neurological disease or disorder that results in atrophy is commonly called a "neurodegenerative disease" or "neurodegenerative disorder." Neurodegenerative diseases and disorders include, but are not limited to, amyotrophic lateral sclerosis (ALS), hereditary spastic hemiplegia, primary lateral sclerosis, spinal muscular atrophy, Kennedy's disease, Alzheimer's disease, Parkinson's disease, multiple sclerosis, and repeat expansion neurodegenerative diseases, *e.g.*, diseases associated with expansions of trinucleotide repeats such as polyglutamine (polyQ) repeat diseases, *e.g.*, Huntington's disease (HD), spinocerebellar ataxia (SCA1, SCA2, SCA3, SCA6, SCA7, and SCA17), spinal and bulbar muscular atrophy (SBMA), dentatorubropallidoluysian atrophy (DRPLA). An example of a disabling neurological disorder that does not appear to result in atrophy is DYT1 dystonia. The gene of interest may encode a ligand for a chemokine involved in the migration of a cancer cell, or a chemokine receptor.

The present invention further provides a method of substantially silencing a target gene of interest or targeted allele for the gene of interest in order to provide a therapeutic effect. As used herein the term "substantially silencing" or "substantially silenced" refers to decreasing, reducing, or inhibiting the expression of the target gene or target allele by at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85% to 100%. As used herein the term "therapeutic effect" refers to a change in the associated abnormalities of the disease state, including pathological and behavioral deficits; a change in the time to progression of the disease state; a reduction, lessening, or alteration of a symptom of the disease; or an improvement in the quality of life of the person afflicted with the disease. Therapeutic effect can be measured quantitatively by a physician or qualitatively by a patient afflicted with the disease state targeted by the siRNA. In certain

embodiments wherein both the mutant and wild type allele are substantially silenced, the term therapeutic effect defines a condition in which silencing of the wild type allele's expression does not have a deleterious or harmful effect on normal functions such that the patient would not have a therapeutic effect.

In one embodiment, the present invention further provides a method of performing allele-specific gene silencing in a mammal by administering to the mammal an isolated first strand of RNA of 15 to 30 nucleotides in length, and an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first strand contains for example at least 15 contiguous nucleotides complementary to a targeted gene of interest, wherein for example at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex for example under physiological conditions, and wherein the siRNA silences only one allele of the targeted gene in the mammal. The alleles of the gene may differ by seven or fewer base pairs, such as by only one base pair. In one example, the gene is a beta-glucuronidase gene. The alleles may be murine-specific and human-specific alleles of beta-glucuronidase. Examples of gene transcripts include an RNA transcript complementary to *TorsinA*, *Ataxin-3*, *huntingtin* or *Tau*. The targeted gene may be a gene associated with a condition amenable to siRNA therapy. For example, the condition amenable to siRNA therapy could be a disabling neurological disorder.

"Neurological disease" and "neurological disorder" refer to both hereditary and sporadic conditions that are characterized by nervous system dysfunction, and which may be associated with atrophy of the affected central or peripheral nervous system structures, or loss of function without atrophy. A neurological disease or disorder that results in atrophy is commonly called a "neurodegenerative disease" or "neurodegenerative disorder." Neurodegenerative diseases and disorders include, but are not limited to, amyotrophic lateral sclerosis (ALS), hereditary spastic hemiplegia, primary lateral sclerosis, spinal muscular atrophy, Kennedy's disease, Alzheimer's disease, Parkinson's disease, multiple sclerosis, and repeat expansion

neurodegenerative diseases, *e.g.*, diseases associated with expansions of trinucleotide repeats such as polyglutamine (polyQ) repeat diseases, *e.g.*, Huntington's disease (HD), spinocerebellar ataxia (SCA1, SCA2, SCA3, SCA6, SCA7, and SCA17), spinal and bulbar muscular atrophy (SBMA), dentatorubropallidoluysian atrophy (DRPLA). An example of a disabling neurological disorder that does not appear to result in atrophy is DYT1 dystonia. The gene of interest may encode a ligand for a chemokine involved in the migration of a cancer cell, or a chemokine receptor.

In one embodiment, the present invention further provides a method of substantially silencing both alleles (*e.g.*, both mutant and wild type alleles) of a target gene. In certain embodiments, the targeting of both alleles of a gene target of interest can confer a therapeutic effect by allowing a certain level of continued expression of the wild-type allele while at the same time inhibiting expression of the mutant (*e.g.*, disease associated) allele at a level that provides a therapeutic effect. For example, a therapeutic effect can be achieved by conferring on the cell the ability to express siRNA as an expression cassette, wherein the expression cassette contains a nucleic acid encoding a small interfering RNA molecule (siRNA) targeted against both alleles, and wherein the expression of the targeted alleles are silenced at a level that inhibits, reduces, or prevents the deleterious gain of function conferred by the mutant allele, but that still allows for adequate expression of the wild type allele at a level that maintains the function of the wild type allele. Examples of such wild type and mutant alleles include without limitation those associated with polyglutamine diseases such as Huntington's Disease.

In one embodiment, the present invention further provides a method of substantially silencing a target allele while allowing expression of a wild-type allele by conferring on the cell the ability to express siRNA as an expression cassette, wherein the expression cassette contains a nucleic acid encoding a small interfering RNA molecule (siRNA) targeted against a target allele, wherein expression from the targeted allele is substantially silenced but wherein expression of the wild-type allele is not substantially silenced.

In one embodiment, the present invention provides a method of treating a dominantly inherited disease in an allele-specific manner by administering to a patient in need thereof an expression cassette, wherein the expression cassette contains a nucleic acid encoding a small interfering RNA molecule (siRNA) targeted against a target allele, wherein expression from the target allele is substantially silenced but wherein expression of the wild-type allele is not substantially silenced.

In one embodiment, the present invention provides a method of treating a dominantly inherited disease by administering to a patient in need thereof an expression cassette, wherein the expression cassette contains a nucleic acid encoding a small interfering RNA molecule (siRNA) targeted against both the mutant allele and the wild type allele of the target gene, wherein expression from the mutant allele is substantially silenced at a level that still allows for expression from the wild type allele to maintain its function in the patient.

In one embodiment, the present invention also provides a method of performing allele-specific gene silencing by administering an expression cassette containing a pol II promoter operably-linked to a nucleic acid encoding at least one strand of a small interfering RNA molecule (siRNA) targeted against a gene of interest, wherein the siRNA silences only one allele of a gene.

In one embodiment, the present invention also provides a method of performing gene silencing by administering an expression cassette containing a pol II promoter operably-linked to a nucleic acid encoding at least one strand of a small interfering RNA molecule (siRNA) targeted against a gene of interest, wherein the siRNA silences one or both alleles of the gene.

In one embodiment, the present invention provides a method of performing allele-specific gene silencing in a mammal by administering to the mammal a vector containing an expression cassette, wherein the expression cassette contains a nucleic acid encoding at least one strand of a small interfering RNA molecule (siRNA) targeted against a gene of interest, wherein the siRNA silences only one allele of a gene.

In one embodiment, the present invention provides a method of performing gene silencing in a mammal by administering to the mammal a vector containing an expression cassette, wherein the expression cassette contains a nucleic acid encoding at least one strand of a small interfering RNA molecule (siRNA) targeted against a gene of interest, wherein the siRNA silences one or both alleles of the gene.

In one embodiment, the present invention provides a method of screening of allele-specific siRNA duplexes, involving contacting a cell containing a predetermined mutant allele with an siRNA with a known sequence, contacting a cell containing a wild-type allele with an siRNA with a known sequence, and determining if the mutant allele is substantially silenced while the wild-type allele retains substantially normal activity.

In one embodiment, the present invention provides a method of screening of specific siRNA duplexes, involving contacting a cell containing both a predetermined mutant allele and a predetermined wild-type allele with an siRNA with a known sequence, and determining if the mutant allele is substantially silenced at a level that allows the wild-type allele to retain substantially normal activity.

In one embodiment, the present invention also provides a method of screening of allele-specific siRNA duplexes involving contacting a cell containing a predetermined mutant allele and a wild-type allele with an siRNA with a known sequence, and determining if the mutant allele is substantially silenced while the wild-type allele retains substantially normal activity.

In one embodiment, the present invention also provides a method for determining the function of an allele by contacting a cell containing a predetermined allele with an siRNA with a known sequence, and determining if the function of the allele is substantially modified.

In one embodiment, the present invention further provides a method for determining the function of an allele by contacting a cell containing a predetermined mutant allele and a wild-type allele with an siRNA with a known

sequence, and determining if the function of the allele is substantially modified while the wild-type allele retains substantially normal function.

In one embodiment, the invention features a method for treating or preventing Huntington's Disease in a subject or organism comprising contacting the subject or organism with a siRNA of the invention under conditions suitable to modulate the expression of the HD gene in the subject or organism whereby the treatment or prevention of Huntington's Disease can be achieved. In one embodiment, the HD gene target comprises a mutant HD allele (*e.g.*, an allele comprising a trinucleotide (CAG) repeat expansion). In one embodiment, the HD gene target comprises both HD allele (*e.g.*, an allele comprising a trinucleotide (CAG) repeat expansion and a wild type allele). The siRNA molecule of the invention can be expressed from vectors as described herein or otherwise known in the art to target appropriate tissues or cells in the subject or organism.

In one embodiment, the invention features a method for treating or preventing Huntington's Disease in a subject or organism comprising, contacting the subject or organism with a siRNA molecule of the invention via local administration to relevant tissues or cells, such as brain cells and tissues (*e.g.*, basal ganglia, striatum, or cortex), for example, by administration of vectors or expression cassettes of the invention that provide siRNA molecules of the invention to relevant cells (*e.g.*, basal ganglia, striatum, or cortex). In one embodiment, the siRNA, vector, or expression cassette is administered to the subject or organism by stereotactic or convection enhanced delivery to the brain. For example, US Patent No. 5,720,720 provides methods and devices useful for stereotactic and convection enhanced delivery of reagents to the brain. Such methods and devices can be readily used for the delivery of siRNAs, vectors, or expression cassettes of the invention to a subject or organism, and is incorporated by reference herein in its entirety. US Patent Application Nos. 2002/0141980; 2002/0114780; and 2002/0187127 all provide methods and devices useful for stereotactic and convection enhanced delivery of reagents that can be readily adapted for delivery of siRNAs, vectors, or expression cassettes of

the invention to a subject or organism, and are incorporated by reference herein in their entirety. Particular devices that may be useful in delivering siRNAs, vectors, or expression cassettes of the invention to a subject or organism are for example described in US Patent Application No. 2004/0162255, which is incorporated by reference herein in its entirety. The siRNA molecule of the invention can be expressed from vectors as described herein or otherwise known in the art to target appropriate tissues or cells in the subject or organism.

In one embodiment, a viral vector of the invention is an AAV vector. An "AAV" vector refers to an adeno-associated virus, and may be used to refer to the naturally occurring wild-type virus itself or derivatives thereof. The term covers all subtypes, serotypes and pseudotypes, and both naturally occurring and recombinant forms, except where required otherwise. As used herein, the term "serotype" refers to an AAV which is identified by and distinguished from other AAVs based on capsid protein reactivity with defined antisera, *e.g.*, there are eight known serotypes of primate AAVs, AAV-1 to AAV-8. For example, serotype AAV-2 is used to refer to an AAV which contains capsid proteins encoded from the cap gene of AAV-2 and a genome containing 5' and 3' ITR sequences from the same AAV-2 serotype. Pseudotyped AAV refers to an AAV that contains capsid proteins from one serotype and a viral genome including 5'-3' ITRs of a second serotype. Pseudotyped rAAV would be expected to have cell surface binding properties of the capsid serotype and genetic properties consistent with the ITR serotype. Pseudotyped rAAV are produced using standard techniques described in the art. As used herein, for example, rAAV1 may be used to refer an AAV having both capsid proteins and 5'-3' ITRs from the same serotype or it may refer to an AAV having capsid proteins from serotype 1 and 5'-3' ITRs from a different AAV serotype, *e.g.*, AAV serotype 2. For each example illustrated herein the description of the vector design and production describes the serotype of the capsid and 5'-3' ITR sequences. The abbreviation "rAAV" refers to recombinant adeno-associated virus, also referred to as a recombinant AAV vector (or "rAAV vector").

An "AAV virus" or "AAV viral particle" refers to a viral particle composed of at least one AAV capsid protein (preferably by all of the capsid proteins of a wild-type AAV) and an encapsidated polynucleotide. If the particle comprises heterologous polynucleotide (*i.e.*, a polynucleotide other than a wild-type AAV genome such as a transgene to be delivered to a mammalian cell), it is typically referred to as "rAAV".

In one embodiment, the AAV expression vectors are constructed using known techniques to at least provide as operatively linked components in the direction of transcription, control elements including a transcriptional initiation region, the DNA of interest and a transcriptional termination region. The control elements are selected to be functional in a mammalian cell. The resulting construct which contains the operatively linked components is flanked (5' and 3') with functional AAV ITR sequences.

By "adeno-associated virus inverted terminal repeats" or "AAV ITRs" is meant the art-recognized regions found at each end of the AAV genome which function together in cis as origins of DNA replication and as packaging signals for the virus. AAV ITRs, together with the AAV rep coding region, provide for the efficient excision and rescue from, and integration of a nucleotide sequence interposed between two flanking ITRs into a mammalian cell genome.

The nucleotide sequences of AAV ITR regions are known. See for example Kotin, R. M. (1994) Human Gene Therapy 5:793-801; Berns, K. I. "Parvoviridae and their Replication" in Fundamental Virology, 2nd Edition, (B. N. Fields and D. M. Knipe, eds.). As used herein, an "AAV ITR" need not have the wild-type nucleotide sequence depicted, but may be altered, *e.g.*, by the insertion, deletion or substitution of nucleotides. Additionally, the AAV ITR may be derived from any of several AAV serotypes, including without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAVX7, etc. Furthermore, 5' and 3' ITRs which flank a selected nucleotide sequence in an AAV vector need not necessarily be identical or derived from the same AAV serotype or isolate, so long as they function as intended, *i.e.*, to allow for excision and rescue of the sequence of interest from a host cell genome or

vector, and to allow integration of the heterologous sequence into the recipient cell genome when AAV Rep gene products are present in the cell.

In one embodiment, AAV ITRs can be derived from any of several AAV serotypes, including without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAVX7, etc. Furthermore, 5' and 3' ITRs which flank a selected nucleotide sequence in an AAV expression vector need not necessarily be identical or derived from the same AAV serotype or isolate, so long as they function as intended, *i.e.*, to allow for excision and rescue of the sequence of interest from a host cell genome or vector, and to allow integration of the DNA molecule into the recipient cell genome when AAV Rep gene products are present in the cell.

In one embodiment, AAV capsids can be derived from any of several AAV serotypes, including without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV6, orAAV8, and the AAV ITRS are derived form AAV serotype 2. Suitable DNA molecules for use in AAV vectors will be less than about 5 kilobases (kb),less than about 4.5 kb, less than about 4kb, less than about 3.5 kb, less than about 3 kb, less than about 2.5 kb in size and are known in the art Dong, J.-Y. et al. (November 10, 1996). "Quantitative Analysis of the Packaging Capacity of Recombinant Adeno-Associated Virus," *Human Gene Ther.* 7(17):2101-2112 and US Patent No. 6,596,535 herein incorporated in its entirety. In some embodiments of the invention the DNA molecules for use in the AAV vectors will contain multiple copies of the identical siRNA sequence. As used herein the term multiple copies of an siRNA sequences means at least 2 copies, at least 3 copies, at least 4 copies, at least 5 copies, at least 6 copies, at least 7 copies, at least 8 copies, at least 9 copies, and at least 10 copies. In some embodiments the DNA molecules for use in the AAV vectors will contain multiple siRNA sequences. As used herein the term multiple = Si RNA sequences means at least 2 siRNA sequences, at least 3 siRNA sequences, at least 4 siRNA sequences, at least 5 siRNA sequences, at least 6 siRNA sequences, at least 7 siRNA sequences, at least 8 siRNA sequences, at least 9 siRNA sequences, and at least 10 siRNA sequences. In some embodiments suitable DNA vectors of the invention will contain a sequence encoding the

siRNA molecule of the invention and a stuffer fragment. Suitable stuffer fragments of the invention include sequences known in the art including without limitation sequences which do not encode an expressed protein molecule; sequences which encode a normal cellular protein which would not have deleterious effect on the cell types in which it was expressed; and sequences which would not themselves encode a functional siRNA duplex molecule.

In one embodiment, suitable DNA molecules for use in AAV vectors will be less than about 5 kilobases (kb) in size and will include, for example, a stuffer sequence and a sequence encoding a siRNA molecule of the invention. For example, in order to prevent any packaging of AAV genomic sequences containing the rep and cap genes, a plasmid containing the rep and cap DNA fragment may be modified by the inclusion of a stuffer fragment as is known in the art into the AAV genome which causes the DNA to exceed the length for optimal packaging. Thus, the helper fragment is not packaged into AAV virions. This is a safety feature, ensuring that only a recombinant AAV vector genome that does not exceed optimal packaging size is packaged into virions. An AAV helper fragment that incorporates a stuffer sequence can exceed the wild-type genome length of 4.6 kb, and lengths above 105% of the wild-type will generally not be packaged. The stuffer fragment can be derived from, for example, such non-viral sources as the Lac-Z or beta-galactosidase gene.

In one embodiment, the selected nucleotide sequence is operably linked to control elements that direct the transcription or expression thereof in the subject in vivo. Such control elements can comprise control sequences normally associated with the selected gene. Alternatively, heterologous control sequences can be employed. Useful heterologous control sequences generally include those derived from sequences encoding mammalian or viral genes. Examples include, but are not limited to, the SV40 early promoter, mouse mammary tumor virus LTR promoter; adenovirus major late promoter (Ad MLP); a herpes simplex virus (HSV) promoter, a cytomegalovirus (CMV) promoter such as the CMV immediate early promoter region (CMVIE), a rous sarcoma virus (RSV) promoter, pol II promoters, pol III promoters, synthetic promoters, hybrid

promoters, and the like. In addition, sequences derived from nonviral genes, such as the murine metallothionein gene, will also find use herein. Such promoter sequences are commercially available from, *e.g.*, Stratagene (San Diego, Calif.).

In one embodiment, both heterologous promoters and other control elements, such as CNS-specific and inducible promoters, enhancers and the like, will be of particular use. Examples of heterologous promoters include the CMB promoter. Examples of CNS-specific promoters include those isolated from the genes from myelin basic protein (MBP), glial fibrillary acid protein (GFAP), and neuron specific enolase (NSE). Examples of inducible promoters include DNA responsive elements for ecdysone, tetracycline, hypoxia and aufin.

In one embodiment, the AAV expression vector which harbors the DNA molecule of interest bounded by AAV ITRs, can be constructed by directly inserting the selected sequence(s) into an AAV genome which has had the major AAV open reading frames ("ORFs") excised therefrom. Other portions of the AAV genome can also be deleted, so long as a sufficient portion of the ITRs remain to allow for replication and packaging functions. Such constructs can be designed using techniques well known in the art. See, *e.g.*, U.S. Pat. Nos. 5,173,414 and 5,139,941; International Publication Nos. WO 92/01070 (published Jan. 23, 1992) and WO 93/03769 (published Mar. 4 1993); Lebkowski *et al.* (1988) Molec. Cell. Biol. 8:3988-3996; Vincent *et al.* (1990) Vaccines 90 (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) Current Opinion in Biotechnology 3:533-539; Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129; Kotin, R. M. (1994) Human Gene Therapy 5:793-801; Shelling and Smith (1994) Gene Therapy 1:165-169; and Zhou *et al.* (1994) J. Exp. Med. 179:1867-1875.

Alternatively, AAV ITRs can be excised from the viral genome or from an AAV vector containing the same and fused 5' and 3' of a selected nucleic acid construct that is present in another vector using standard ligation techniques, such as those described in Sambrook *et al.*, supra. For example, ligations can be accomplished in 20 mM Tris-Cl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 33 μg/ml BSA, 10 mM-50 mM NaCl, and either 40 uM ATP, 0.01-0.02 (Weiss) units T4

DNA ligase at 0°C. (for "sticky end" ligation) or 1 mM ATP, 0.3-0.6 (Weiss) units T4 DNA ligase at 14°C. (for "blunt end" ligation). Intermolecular "sticky end" ligations are usually performed at 30-100 μg/ml total DNA concentrations (5-100 nM total end concentration). AAV vectors which contain ITRs have been described in, *e.g.*, U.S. Pat. No. 5,139,941. In particular, several AAV vectors are described therein which are available from the American Type Culture Collection ("ATCC") under Accession Numbers 53222, 53223, 53224, 53225 and 53226.

Additionally, chimeric genes can be produced synthetically to include AAV ITR sequences arranged 5' and 3' of one or more selected nucleic acid sequences. Preferred codons for expression of the chimeric gene sequence in mammalian CNS cells can be used. The complete chimeric sequence is assembled from overlapping oligonucleotides prepared by standard methods. See, *e.g.*, Edge, Nature (1981) 292:756; Nambair *et al.* Science (1984) 223:1299; Jay *et al.* J. Biol. Chem. (1984) 259:6311.

In order to produce rAAV virions, an AAV expression vector is introduced into a suitable host cell using known techniques, such as by transfection. A number of transfection techniques are generally known in the art. See, *e.g.*, Graham *et al.* (1973) Virology, 52:456, Sambrook *et al.* (1989) Molecular Cloning, a laboratory manual, Cold Spring Harbor Laboratories, New York, Davis *et al.* (1986) Basic Methods in Molecular Biology, Elsevier, and Chu *et al.* (1981) Gene 13:197. Particularly suitable transfection methods include calcium phosphate co-precipitation (Graham *et al.* (1973) Virol. 52:456-467), direct micro-injection into cultured cells (Capecchi, M. R. (1980) Cell 22:479-488), electroporation (Shigekawa *et al.* (1988) BioTechniques 6:742-751), liposome mediated gene transfer (Mannino *et al.* (1988) BioTechniques 6:682-690), lipid-mediated transduction (Felgner *et al.* (1987) Proc. Natl. Acad. Sci. USA 84:7413-7417), and nucleic acid delivery using high-velocity microprojectiles (Klein *et al.* (1987) Nature 327:70-73).

In one embodiment, suitable host cells for producing rAAV virions include microorganisms, yeast cells, insect cells, and mammalian cells, that can

be, or have been, used as recipients of a heterologous DNA molecule. The term includes the progeny of the original cell which has been transfected. Thus, a "host cell" as used herein generally refers to a cell which has been transfected with an exogenous DNA sequence. Cells from the stable human cell line, 293 (readily available through, *e.g.*, the American Type Culture Collection under Accession Number ATCC CRL1573) can be used in the practice of the present invention. Particularly, the human cell line 293 is a human embryonic kidney cell line that has been transformed with adenovirus type-5 DNA fragments (Graham *et al.* (1977) J. Gen. Virol. 36:59), and expresses the adenoviral E1a and E1b genes (Aiello *et al.* (1979) Virology 94:460). The 293 cell line is readily transfected, and provides a particularly convenient platform in which to produce rAAV virions.

In one embodiment, host cells containing the above-described AAV expression vectors are rendered capable of providing AAV helper functions in order to replicate and encapsidate the nucleotide sequences flanked by the AAV ITRs to produce rAAV virions. AAV helper functions are generally AAV-derived coding sequences which can be expressed to provide AAV gene products that, in turn, function in trans for productive AAV replication. AAV helper functions are used herein to complement necessary AAV functions that are missing from the AAV expression vectors. Thus, AAV helper functions include one, or both of the major AAV ORFs, namely the rep and cap coding regions, or functional homologues thereof.

The Rep expression products have been shown to possess many functions, including, among others: recognition, binding and nicking of the AAV origin of DNA replication; DNA helicase activity; and modulation of transcription from AAV (or other heterologous) promoters. The Cap expression products supply necessary packaging functions. AAV helper functions are used herein to complement AAV functions in trans that are missing from AAV vectors.

The term "AAV helper construct" refers generally to a nucleic acid molecule that includes nucleotide sequences providing AAV functions deleted

from an AAV vector which is to be used to produce a transducing vector for delivery of a nucleotide sequence of interest. AAV helper constructs are commonly used to provide transient expression of AAV rep and/or cap genes to complement missing AAV functions that are necessary for lytic AAV replication; however, helper constructs lack AAV ITRs and can neither replicate nor package themselves. AAV helper constructs can be in the form of a plasmid, phage, transposon, cosmid, virus, or virion. A number of AAV helper constructs have been described, such as the commonly used plasmids pAAV/Ad and pIM29+45 which encode both Rep and Cap expression products. See, *e.g.*, Samulski *et al.* (1989) J. Virol. 63:3822-3828; and McCarty *et al.* (1991) J. Virol. 65:2936-2945. A number of other vectors have been described which encode Rep and/or Cap expression products. See, *e.g.*, U.S. Pat. No. 5,139,941.

By "AAV rep coding region" is meant the art-recognized region of the AAV genome which encodes the replication proteins Rep 78, Rep 68, Rep 52 and Rep 40. These Rep expression products have been shown to possess many functions, including recognition, binding and nicking of the AAV origin of DNA replication, DNA helicase activity and modulation of transcription from AAV (or other heterologous) promoters. The Rep expression products are collectively required for replicating the AAV genome. For a description of the AAV rep coding region, see, *e.g.*, Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129; and Kotin, R. M. (1994) Human Gene Therapy 5:793-801. Suitable homologues of the AAV rep coding region include the human herpesvirus 6 (HHV-6) rep gene which is also known to mediate AAV-2 DNA replication (Thomson *et al.* (1994) Virology 204:304-311).

By "AAV cap coding region" is meant the art-recognized region of the AAV genome which encodes the capsid proteins VP1, VP2, and VP3, or functional homologues thereof. These Cap expression products supply the packaging functions which are collectively required for packaging the viral genome. For a description of the AAV cap coding region, see, *e.g.*, Muzyczka, N. and Kotin, R. M. (supra).

In one embodiment, AAV helper functions are introduced into the host cell by transfecting the host cell with an AAV helper construct either prior to, or concurrently with, the transfection of the AAV expression vector. AAV helper constructs are thus used to provide at least transient expression of AAV rep and/or cap genes to complement missing AAV functions that are necessary for productive AAV infection. AAV helper constructs lack AAV ITRs and can neither replicate nor package themselves. These constructs can be in the form of a plasmid, phage, transposon, cosmid, virus, or virion. A number of AAV helper constructs have been described, such as the commonly used plasmids pAAV/Ad and pIM29+45 which encode both Rep and Cap expression products. See, *e.g.*, Samulski *et al.* (1989) J. Virol. 63:3822-3828; and McCarty *et al.* (1991) J. Virol. 65:2936-2945. A number of other vectors have been described which encode Rep and/or Cap expression products. See, *e.g.*, U.S. Pat. No. 5,139,941.

In one embodiment, both AAV expression vectors and AAV helper constructs can be constructed to contain one or more optional selectable markers. Suitable markers include genes which confer antibiotic resistance or sensitivity to, impart color to, or change the antigenic characteristics of those cells which have been transfected with a nucleic acid construct containing the selectable marker when the cells are grown in an appropriate selective medium. Several selectable marker genes that are useful in the practice of the invention include the hygromycin B resistance gene (encoding Aminoglycoside phosphotranferase (APH)) that allows selection in mammalian cells by conferring resistance to G418 (available from Sigma, St. Louis, Mo.). Other suitable markers are known to those of skill in the art.

In one embodiment, the host cell (or packaging cell) is rendered capable of providing non AAV derived functions, or "accessory functions," in order to produce rAAV virions. Accessory functions are non AAV derived viral and/or cellular functions upon which AAV is dependent for its replication. Thus, accessory functions include at least those non AAV proteins and RNAs that are required in AAV replication, including those involved in activation of AAV gene transcription, stage specific AAV mRNA splicing, AAV DNA replication,

synthesis of Cap expression products and AAV capsid assembly. Viral-based accessory functions can be derived from any of the known helper viruses.

In one embodiment, accessory functions can be introduced into and then expressed in host cells using methods known to those of skill in the art. Commonly, accessory functions are provided by infection of the host cells with an unrelated helper virus. A number of suitable helper viruses are known, including adenoviruses; herpesviruses such as herpes simplex virus types 1 and 2; and vaccinia viruses. Nonviral accessory functions will also find use herein, such as those provided by cell synchronization using any of various known agents. See, *e.g.*, Buller *et al.* (1981) J. Virol. 40:241-247; McPherson *et al.* (1985) Virology 147:217-222; Schlehofer *et al.* (1986) Virology 152:110-117.

In one embodiment, accessory functions are provided using an accessory function vector. Accessory function vectors include nucleotide sequences that provide one or more accessory functions. An accessory function vector is capable of being introduced into a suitable host cell in order to support efficient AAV virion production in the host cell. Accessory function vectors can be in the form of a plasmid, phage, transposon or cosmid. Accessory vectors can also be in the form of one or more linearized DNA or RNA fragments which, when associated with the appropriate control elements and enzymes, can be transcribed or expressed in a host cell to provide accessory functions. See, for example, International Publication No. WO 97/17548, published May 15, 1997.

In one embodiment, nucleic acid sequences providing the accessory functions can be obtained from natural sources, such as from the genome of an adenovirus particle, or constructed using recombinant or synthetic methods known in the art. In this regard, adenovirus-derived accessory functions have been widely studied, and a number of adenovirus genes involved in accessory functions have been identified and partially characterized. See, *e.g.*, Carter, B. J. (1990) "Adeno-Associated Virus Helper Functions," in CRC Handbook of Parvoviruses, vol. I (P. Tijssen, ed.), and Muzyczka, N. (1992) Curr. Topics. Microbiol and Immun. 158:97-129. Specifically, early adenoviral gene regions E1 a, E2a, E4, VAI RNA and, possibly, E1b are thought to participate in the

accessory process. Janik *et al.* (1981) Proc. Natl. Acad. Sci. USA 78:1925-1929. Herpesvirus-derived accessory functions have been described. See, *e.g.*, Young *et al.* (1979) Prog. Med. Virol. 25:113. Vaccinia virus-derived accessory functions have also been described. See, *e.g.*, Carter, B. J. (1990), supra., Schlehofer *et al.* (1986) Virology 152:110-117.

In one embodiment, as a consequence of the infection of the host cell with a helper virus, or transfection of the host cell with an accessory function vector, accessory functions are expressed which transactivate the AAV helper construct to produce AAV Rep and/or Cap proteins. The Rep expression products excise the recombinant DNA (including the DNA of interest) from the AAV expression vector. The Rep proteins also serve to duplicate the AAV genome. The expressed Cap proteins assemble into capsids, and the recombinant AAV genome is packaged into the capsids. Thus, productive AAV replication ensues, and the DNA is packaged into rAAV virions.

In one embodiment, following recombinant AAV replication, rAAV virions can be purified from the host cell using a variety of conventional purification methods, such as CsCl gradients. Further, if infection is employed to express the accessory functions, residual helper virus can be inactivated, using known methods. For example, adenovirus can be inactivated by heating to temperatures of approximately 60.degrees C. for, *e.g.*, 20 minutes or more. This treatment effectively inactivates only the helper virus since AAV is extremely heat stable while the helper adenovirus is heat labile. The resulting rAAV virions are then ready for use for DNA delivery to the CNS (*e.g.*, cranial cavity) of the subject.

Methods of delivery of viral vectors include, but are not limited to, intra-arterial, intra-muscular, intravenous, intranasal and oral routes. Generally, rAAV virions may be introduced into cells of the CNS using either in vivo or in vitro transduction techniques. If transduced in vitro, the desired recipient cell will be removed from the subject, transduced with rAAV virions and reintroduced into the subject. Alternatively, syngeneic or xenogeneic cells can be used where those cells will not generate an inappropriate immune response in the subject.

Suitable methods for the delivery and introduction of transduced cells into a subject have been described. For example, cells can be transduced in vitro by combining recombinant AAV virions with CNS cells *e.g.*, in appropriate media, and screening for those cells harboring the DNA of interest can be screened using conventional techniques such as Southern blots and/or PCR, or by using selectable markers. Transduced cells can then be formulated into pharmaceutical compositions, described more fully below, and the composition introduced into the subject by various techniques, such as by grafting, intramuscular, intravenous, subcutaneous and intraperitoneal injection.

In one embodiment, for in vivo delivery, the rAAV virions are formulated into pharmaceutical compositions and will generally be administered parenterally, *e.g.*, by intramuscular injection directly into skeletal or cardiac muscle or by injection into the CNS.

In one embodiment, viral vectors of the invention are delivered to the CNS via convection-enhanced delivery (CED) systems that can efficiently deliver viral vectors, *e.g.*, AAV, over large regions of a subject's brain (*e.g.*, striatum and/or cortex). As described in detail and exemplified below, these methods are suitable for a variety of viral vectors, for instance AAV vectors carrying therapeutic genes (*e.g.*, siRNAs).

Any convection-enhanced delivery device may be appropriate for delivery of viral vectors. In one embodiment, the device is an osmotic pump or an infusion pump. Both osmotic and infusion pumps are commerically available from a variety of suppliers, for example Alzet Corporation, Hamilton Corporation, Aiza, Inc., Palo Alto, Calif.). Typically, a viral vector is delivered via CED devices as follows. A catheter, cannula or other injection device is inserted into CNS tissue in the chosen subject. In view of the teachings herein, one of skill in the art could readily determine which general area of the CNS is an appropriate target. For example, when delivering AAV vector encoding a therapeutic gene to treat PD, the striatum is a suitable area of the brain to target. Stereotactic maps and positioning devices are available, for example from ASI Instruments, Warren, Mich. Positioning may also be conducted by using

anatomical maps obtained by CT and/or MRI imaging of the subject's brain to help guide the injection device to the chosen target. Moreover, because the methods described herein can be practiced such that relatively large areas of the brain take up the viral vectors, fewer infusion cannula are needed. Since surgical complications are related to the number of penetrations, the methods described herein also serve to reduce the side effects seen with conventional delivery techniques.

In one embodiment, pharmaceutical compositions will comprise sufficient genetic material to produce a therapeutically effective amount of the siRNA of interest, *i.e.*, an amount sufficient to reduce or ameliorate symptoms of the disease state in question or an amount sufficient to confer the desired benefit. The pharmaceutical compositions will also contain a pharmaceutically acceptable excipient. Such excipients include any pharmaceutical agent that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Pharmaceutically acceptable excipients include, but are not limited to, sorbitol, Tween80, and liquids such as water, saline, glycerol and ethanol. Pharmaceutically acceptable salts can be included therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. A thorough discussion of pharmaceutically acceptable excipients is available in REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub. Co., N.J. 1991).

As is apparent to those skilled in the art in view of the teachings of this specification, an effective amount of viral vector which must be added can be empirically determined. Administration can be effected in one dose, continuously or intermittently throughout the course of treatment. Methods of determining the most effective means and dosages of administration are well known to those of skill in the art and will vary with the viral vector, the

composition of the therapy, the target cells, and the subject being treated. Single and multiple administrations can be carried out with the dose level and pattern being selected by the treating physician.

It should be understood that more than one transgene could be expressed by the delivered viral vector. Alternatively, separate vectors, each expressing one or more different transgenes, can also be delivered to the CNS as described herein. Furthermore, it is also intended that the viral vectors delivered by the methods of the present invention be combined with other suitable compositions and therapies.

Brief Description of the Figures

Figure 1. siRNA expressed from CMV promoter constructs and *in vitro* effects. (A) A cartoon of the expression plasmid used for expression of functional siRNA in cells. The CMV promoter was modified to allow close juxtaposition of the hairpin to the transcription initiation site, and a minimal polyadenylation signal containing cassette was constructed immediately 3' of the MCS (mCMV, modified CMV; mpA, minipA). (B, C) Fluorescence photomicrographs of HEK293 cells 72 h after transfection of pEGFPN1 and pCMVβgal (control), or pEGFPN1 and pmCMVsiGFPmpA, respectively. (D) Northern blot evaluation of transcripts harvested from pmCMVsiGFPmpA (lanes 3, 4) and pmCMVsißgalmpA (lane 2) transfected HEK293 cells. Blots were probed with ^{32}P-labeled sense oligonucleotides. Antisense probes yielded similar results (not shown). Lane 1, ^{32}P-labeled RNA markers. AdsiGFP infected cells also possessed appropriately sized transcripts (not shown). (E) Northern blot for evaluation of target mRNA reduction by siRNA (upper panel). The internal control GAPDH is shown in the lower panel. HEK293 cells were transfected with pEGFPN1 and pmCMVsiGFPmpA, expressing siGFP, or

plasmids expressing the control siRNA as indicated. pCMVeGFPx, which expresses siGFPx, contains a large poly(A) cassette from SV40 large T and an unmodified CMV promoter, in contrast to pmCMVsiGFPmpA shown in (A). (F) Western blot with anti-GFP antibodies of cell lysates harvested 72 h after transfection with pEGFPN1 and pCMVsiGFPmpA, or pEGFPN1 and pmCMVsiβglucmpA. (G, H) Fluorescence photomicrographs of HEK293 cells 72 h after transfection of pEGFPN1 and pCMVsiGFPx, or pEGFPN1 and pmCMVsißglucmpA, respectively. (I, J) siRNA reduces expression from endogenous alleles. Recombinant adenoviruses were generated from pmCMVsiβglucmpA and pmCMVsiGFPmpA and purified. HeLa cells were infected with 25 infectious viruses/cell (MOI = 25) or mock-infected (control) and cell lysates harvested 72 h later. (I) Northern blot for ß-glucuronidase mRNA levels in Adsißgluc and AdsiGFP transduced cells. GAPDH was used as an internal control for loading. (J) The concentration of β-glucuronidase activity in lysates quantified by a fluorometric assay. Stein, C.S. *et al.*, *J. Virol.* **73**:3424-3429 (1999).

Figure 2. Viral vectors expressing siRNA reduce expression from transgenic and endogenous alleles *in vivo*. Recombinant adenovirus vectors were prepared from the siGFP and siβgluc shuttle plasmids described in Fig. 1. (A) Fluorescence microscopy reveals diminution of eGFP expression *in vivo*. In addition to the siRNA sequences in the E1 region of adenovirus, RFP expression cassettes in E3 facilitate localization of gene transfer. Representative photomicrographs of eGFP (left), RFP (middle), and merged images (right) of coronal sections from mice injected with adenoviruses expressing siGFP (top panels) or siβgluc (bottom panels) demonstrate siRNA specificity in eGFP transgenic mice striata after direct brain injection. (B) Full coronal brain sections (1 mm) harvested from AdsiGFP or Adsiβgluc injected mice were split into hemisections and both ipsilateral (il) and contralateral (cl) portions evaluated by western blot using antibodies to GFP. Actin was used as an internal control for each sample. (C) Tail vein injection of recombinant

adenoviruses expressing siβgluc directed against mouse β-glucuronidase (AdsiMuβgluc) reduces endogenous β–glucuronidase RNA as determined by Northern blot in contrast to control-treated (Adsiβgal) mice.

Figure 3. siGFP gene transfer reduces Q19-eGFP expression in cell lines. PC12 cells expressing the polyglutamine repeat Q19 fused to eGFP (eGFP-Q19) under tetracycline repression (A, bottom left) were washed and dox-free media added to allow eGFP-Q19 expression (A, top left). Adenoviruses were applied at the indicated multiplicity of infection (MOI) 3 days after dox removal. (A) eGFP fluorescence 3 days after adenovirus-mediated gene transfer of Adsiβgluc (top panels) or AdsiGFP (bottom panels). (B, C) Western blot analysis of cell lysates harvested 3 days after infection at the indicated MOIs demonstrate a dose-dependent decrease in GFP-Q19 protein levels. NV, no virus. Top lanes, eGFP-Q19. Bottom lanes, actin loading controls. (D) Quantitation of eGFP fluorescence. Data represent mean total area fluorescence ± standard deviation in 4 low power fields/well (3 wells/plate).

Figure 4. siRNA mediated reduction of expanded polyglutamine protein levels and intracellular aggregates. PC12 cells expressing tet-repressible eGFP-Q80 fusion proteins were washed to remove doxycycline and adenovirus vectors expressing siRNA were applied 3 days later. (A-D) Representative punctate eGFP fluorescence of aggregates in mock-infected cells (A), or those infected with 100 MOI of Adsiβgluc (B), AdsiGFPx (C) or Adsiβgal (D). (E) Three days after infection of dox-free eGFP-Q80 PC12 cells with AdsiGFP, aggregate size and number are notably reduced. (F) Western blot analysis of eGFP-Q80 aggregates (arrowhead) and monomer (arrow) following Adsiβgluc or AdsiGFP infection at the indicated MOIs demonstrates dose dependent siGFP-mediated reduction of GFP-Q80 protein levels. (G) Quantification of the total area of fluorescent inclusions measured in 4 independent fields/well 3 days after virus was applied at the indicated MOIs. The data are mean ± standard deviation.

Figure 5. RNAi-mediated suppression of expanded CAG repeat containing genes. Expanded CAG repeats are not direct targets for preferential inactivation (A), but a linked SNP can be exploited to generate siRNA that selectively

silences mutant ataxin-3 expression (B-F). (A) Schematic of cDNA encoding generalized polyQ-fluorescent protein fusions. Bars indicate regions targeted by siRNAs. HeLa cells co-transfected with Q80-GFP, Q19-RFP and the indicated siRNA. Nuclei are visualized by DAPI staining (blue) in merged images. (B)Schematic of human *ataxin-3* cDNA with bars indicating regions targeted by siRNAs. The targeted SNP (G987C) is shown in color. In the displayed siRNAs, red or blue bars denote C or G respectively. In this Figure, AGCAGCAGCAGGGGGACCTATCAGGAC is SEQ ID NO:7, and CAGCAGCAGCAGCGGGACCTATCAGGAC is SEQ ID NO:8. (C) Quantitation of fluorescence in Cos-7 cells transfected with wild type or mutant ataxin-3-GFP expression plasmids and the indicated siRNA. Fluorescence from cells co-transfected with siMiss was set at one. Bars depict mean total fluorescence from three independent experiments +/- standard error of the mean (SEM). (D) Western blot analysis of cells co-transfected with the indicated ataxin-3 expression plasmids (top) and siRNAs (bottom). Appearance of aggregated, mutant ataxin-3 in the stacking gel (seen with siMiss and siG10) is prevented by siRNA inhibition of the mutant allele. (E) Allele specificity is retained in the simulated heterozygous state. Western blot analysis of Cos-7 cells cotransfected with wild-type (atx-3-Q28-GFP) and mutant (atx-Q166) expression plasmids along with the indicated siRNAs. (Mutant ataxin-3 detected with 1C2, an antibody specific for expanded polyQ, and wild-type ataxin-3 detected with anti-ataxin-3 antibody.) (F) Western blot of Cos-7 cells transfected with Atx-3-GFP expression plasmids and plasmids encoding the indicated shRNA. The negative control plasmid, phU6-LacZi, encodes siRNA specific for LacZ. Both normal and mutant protein were detected with anti-ataxin-3 antibody. Tubulin immunostaining shown as a loading control in panels (D)-(F).

Figure 6. Primer sequences for *in vitro* synthesis of siRNAs using T7 polymerase. All primers contain the following T7 promoter sequence at their 3' ends: 5'-TATAGTGAGTCGTATTA-3' (SEQ ID NO:9). The following primer was annealed to all oligos to synthesize siRNAs: 5'-TAATACGACTCACTATAG-3' (SEQ ID NO:10).

Figure 7. Inclusion of either two (siC7/8) or three (siC10) CAG triplets at the 5' end of ataxin-3 siRNA does not inhibit expression of unrelated CAG repeat containing genes. (A) Western blot analysis of Cos-7 cells transfected with CAG repeat-GFP fusion proteins and the indicated siRNA. Immunostaining with monoclonal anti-GFP antibody (MBL) at 1:1000 dilution. (B) Western blot analysis of Cos-7 cells transfected with Flag-tagged ataxin-1-Q30, which is unrelated to ataxin-3, and the indicated siRNA. Immunostaining with anti-Flag monoclonal antibody (Sigma St. Louis, MO) at 1:1000 dilution. In panels (A) and (B), lysates were collected 24 hours after transfection. Tubulin immunostaining shown as a loading control.

Figure 8. shRNA-expressing adenovirus mediates allele-specific silencing in transiently transfected Cos-7 cells simulating the heterozygous state. (A) Representative images of cells cotransfected to express wild type and mutant ataxin-3 and infected with the indicated adenovirus at 50 multiplicities of infection (MOI). Atx-3-Q28-GFP (green) is directly visualized and Atx-3-Q166 (red) is detected by immunofluorescence with 1C2 antibody. Nuclei visualized with DAPI stain in merged images. An average of 73.1% of cells co-expressed both ataxin-3 proteins with siMiss. (B) Quantitation of mean fluorescence from 2 independent experiments performed as in (A). (C) Western blot analysis of viral-mediated silencing in Cos-7 cells expressing wild type and mutant ataxin-3 as in (A). Mutant ataxin-3 detected with 1C2 antibody and wild-type human and endogenous primate ataxin-3 detected with anti-ataxin-3 antibody. (D) shRNA-expressing adenovirus mediates allele-specific silencing in stably transfected neural cell lines. Differentiated PC12 neural cells expressing wild type (left) or mutant (right) ataxin-3 were infected with adenovirus (100 MOI) engineered to express the indicated hairpin siRNA. Shown are Western blots immunostained for ataxin-3 and GAPDH as loading control.

Figure 9. Allele-specific siRNA suppression of a missense Tau mutation. (A) Schematic of human tau cDNA with bars indicating regions and mutations tested for siRNA suppression. Of these, the V337M region showed effective suppression and was further studied. Vertical bars represent

microtubule binding repeat elements in Tau. In the displayed siRNAs, blue and red bars denote A and C respectively. In this Figure, GTGGCCAGATGGAAGTAAAATC is SEQ ID NO:35, and GTGGCCAGGTGGAAGTAAAATC is SEQ ID NO:41. (B) Western blot analysis of cells co-transfected with WT or V337M Tau-EGFP fusion proteins and the indicated siRNAs. Cells were lysed 24 hr after transfection and probed with anti-tau antibody. Tubulin immunostaining is shown as loading control. (C) Quantitation of fluorescence in Cos-7 cells transfected with wild type tau-EGFP or mutant V337M tau-EGFP expression plasmids and the indicated siRNAs. Bars depict mean fluorescence and SEM from three independent experiments. Fluorescence from cells co-transfected with siMiss was set at one.

Figure 10. Allele-specific silencing of Tau in cells simulating the heterozygous state. (A) Representative fluorescent images of fixed Hela cells co-transfected with flag-tagged WT-Tau (red), V337M-Tau-GFP (green), and the indicated siRNAs. An average of 73.7% of cells co-expressed both Tau proteins with siMiss. While siA9 suppresses both alleles, siA9/C12 selectively decreased expression of mutant Tau only. Nuclei visualized with DAPI stain in merged images. (B) Quantitation of mean fluorescence from 2 independent experiments performed as in (A). (C) Western blot analysis of cells co-transfected with Flag-WT-Tau and V337M-Tau-EGFP fusion proteins and the indicated siRNAs. Cells were lysed 24 hr after transfection and probed with anti-tau antibody. V337M-GFP Tau was differentiated based on reduced electrophoretic mobility due to the addition of GFP. Tubulin immunostaining is shown as a loading control.

Figure 11. Schematic diagram of allele-specific silencing of mutant TorsinA by small interfering RNA (siRNA). In the disease state, wild type and mutant alleles of *TOR1A* are both transcribed into mRNA. siRNA with sequence identical to the mutant allele (deleted of GAG) should bind mutant mRNA selectively and mediate its degradation by the RNA-induced silencing complex (RISC) (circle). Wild type mRNA, not recognized by the mutant –specific siRNA, will remain and continue to be translated into normal TorsinA. The two

adjacent GAG's in wild type *TOR1A* alleles are shown as two parallelograms, one of which is deleted in mutant *TOR1A* alleles.

Figure 12. Design and targeted sequences of siRNAs. Shown are the relative positions and targeted mRNA sequences for each primer used in this study. Mis-siRNA (negative control) does not target TA; com-siRNA targets a sequence present in wild type and mutant TA; wt-siRNA targets only wild type TA; and three mutant-specific siRNAs (Mut A, B, C). preferentially target mutant TA. The pair of GAG codons near the c-terminus of wild type mRNA are shown in underlined gray and black, with one codon deleted in mutant mRNA.

Figure 13. siRNA silencing of TAwt and TAmut in Cos-7 cells. (A) Western blot results showing the effect of different siRNAs on GFP-TAwt expression levels. Robust suppression is achieved with wt-siRNA and com-siRNA, while the mutant-specific siRNAs MutA, (B) and (C) have modest or no effect on GFP-TAwt expression. Tubulin loading controls are also shown. (B) Similar experiments with cells expressing HA-TAmut, showing significant suppression by mutant-specific siRNAs and com-siRNA but no suppression by the wild type-specific siRNA, wt-siRNA. (C) Quantification of results from at least three separate experiments as in A and B. (D) Cos-7 cells transfected with GFP-TAwt or GFP-TAmut and different siRNAs visualized under fluorescence microscopy (200X). Representative fields are shown indicating allele-specific suppression. (E) Quantification of fluorescence signal from two different experiments as in D.

Figure 14. Allele-specific silencing by siRNA in the simulated heterozygous state. Cos-7 cells were cotransfected with plasmids encoding differentially tagged TAwt and TAmut, together with the indicated siRNA. (A) Western blot results analysis showing selective suppression of the targeted allele by wt-siRNA or mutC-siRNA. (B) Quantification of results from three experiments as in (A).

Figure 15. Allele-specific silencing of mutant huntingtin by siRNA. PC6-3 cells were co-transfected with plasmids expressing siRNA specific for the polymorphism encoding the transcript for mutant huntingtin.

Figure 16. RNAi reduces human huntingtin expression *in vitro*. *(A)* RNA sequence of shHD2.1. The 21 nucleotide antisense strand is cognate to nucleotides 416-436 of human htt mRNA (Genbank #NM 00211). ***(B and C)*** Northern and western blots demonstrate shHD2.1 mediated reduction of HD-N171-82Q mRNA and protein expression, 48 h post-transfection of target- and shRNA-expressing plasmids. GAPDH and actin serve as loading controls. ***(D)*** Western blots show that shHD2.1 inhibits expression of full-length human huntingtin protein, 48 h post-transfection. ***(E)*** ShHD2.1 induces dose-dependent reduction of human htt mRNA. Cells were transfected with shLacZ- or shHD2.1-expressing plasmids in the indicated amounts. Relative htt expression was determined by quantitative PCR 24 h later. SEQ ID NO:56 is 5′-AAGAAAGAACUUUCAGCUACC-3′. SEQ ID NO:57 is 5′-GGUAGCUGAAAGUUCUUUCUU-3′. SEQ ID NO:58 is 5′-GAAGCUUG-3′. SEQ ID NO:59 is 5′-AAGAAAGAACUUUCAGCUACCGAAGCUUGGGUAGCUGAAAGUUCUUUCUUUUUUUU-3′.

Figure 17. AAV.shHD2.1 delivers widespread RNAi expression to mouse striatum. *(A)* AAV.shHD2.1 viral vector. ITR, inverted terminal repeat. ***(B)*** Northern blot showing shHD2.1 transcripts are expressed *in vivo*. Processed antisense (lower band) and unprocessed (upper band) shHD2.1 transcripts in three different AAV.shHD2.1-injected mice. L, ladder; +, positive control oligo. Blot was probed with radiolabeled sense probe. ***(C)*** Typical AAV1 transduction pattern (hrGFP) in mouse brain. CC, corpus callosum; LV, lateral ventricle.

Figure 18. AAV.shHD2.1 eliminates accumulation of huntingtin-reactive neuronal inclusions and reduces HD-N171-82Q mRNA *in vivo*. ***(A)*** Representative photomicrographs show htt-reactive inclusions (arrows) in HD striatal cells transduced with AAV.shLacZ-, but not AAV.shHD2.1. Scale bar, 20 μm. ***(B)*** Higher magnification photomicrograph from a (bottom, right) showing lack of htt-reactive inclusions in cells transduced by AAV.shHD2.1. * serves as a marker for orientation. Scale bar, 20 μm. ***(C)*** Representative western blot demonstrates decreased HD-N171-82Q expression in mouse striata

transduced with AAV.shHD2.1 compared to uninjected or AAV.shLacZ-injected striata. Prion protein was used as a loading control to normalize for tissues expressing the HD-N171-82Q transgene. *(D)* AAV.shHD2.1-treated HD mice showed a 55% average reduction in HD-N171-82Q mRNA compared to AAV.shLacZ or uninjected HD mice. Data are means + S.E.M. relative to uninjected HD samples. *, difference from AAV.shHD2.1 samples, $p<0.05$ (ANOVA). *(E)* Mice were injected directly into cerebellum with AAV.shHD2.1 or AAV.shLacZ. Cerebellar sections confirm that AAV.shHD2.1, but not AAV.shLacZ, reduces htt immunoreactivity. GCL, granule cell layer; ML, molecular layer. Scale bar, 100 μm.

Figure 19. AAV.shHD2.1 improves behavioral deficits in HD-N171-82Q mice. *(A)* Box plot. Bilateral striatal delivery of AAV.shHD2.1 improves stride length in HD-N171-82Q mice. HD mice had significantly shorter stride lengths compared to WT. AAV.shHD2.1 mediated significant gait improvement relative to control-treated HD mice. *, $p<0.0001$ (ANOVA, Scheffe post-hoc). *(B)* Bilateral striatal delivery of AAV.shHD2.1 significantly improves rotarod performance in HD-N171-82Q mice. Only AAV.shLacZ-injected and uninjected HD-N171-82Q declined significantly with time. Data are means ± S.E.M.

Figure 20. DNA sequences of huntingtin hairpins. The bases that are underlined indicate changes from the native huntingtin sequence.

Figure 21. PCR method for cloning hairpins. A 79 nt primer is used with the Ampr template. Pfu and DMSO are used in the amplification reaction. Products are ligated directly into pCR-Blunt Topo (Invitrogen) and Kanr resistant colonies picked and sequenced. Positive clones can be used directly.

Figure 22. Reduction of eGFP inclusions after transduction with 25, 50 or 100 viruses/cell into cultures with pre-formed aggregates. Note dose-dependent response with shGFP vectors only.

Figure 23. Regulated RNAi. Two Teto2 sequences were placed up- and down-stream of the TATA box of the H1 promoter element (cartoon). Either control shRNA or shGFP was placed into the cassette for expression of hairpins.

Plasmids expressing GFP and the hairpin constructs were transfected into a cell line expressing the TetR (tet-repressor). GFP fluorescence (left panels) or western blot (right panels) was evaluated in the absence (TetR binding) or presence (TetR off) of doxycycline.

Figure 24. Top, FIV construct. Bottom, AAV construct. Both express the hrGFP reporter so that transduced cells can be readily evaluated for shRNA efficacy (as in Figures 3 and 4).

Detailed Description of the Invention

Modulation of gene expression by endogenous, noncoding RNAs is increasingly appreciated as a mechanism playing a role in eukaryotic development, maintenance of chromatin structure and genomic integrity (McManus, 2002). Recently, techniques have been developed to trigger RNA interference (RNAi) against specific targets in mammalian cells by introducing exogenously produced or intracellularly expressed siRNAs (Elbashir, 2001; Brummelkamp, 2002). These methods have proven to be quick, inexpensive and effective for knockdown experiments *in vitro* and *in vivo* (2 Elbashir, 2001; Brummelkamp, 2002; McCaffrey, 2002; Xia, 2002). The ability to accomplish selective gene silencing has led to the hypothesis that siRNAs might be employed to suppress gene expression for therapeutic benefit (Xia, 2002; Jacque, 2002; Gitlin, 2002).

RNA interference is now established as an important biological strategy for gene silencing, but its application to mammalian cells has been limited by nonspecific inhibitory effects of long double-stranded RNA on translation. Moreover, delivery of interfering RNA has largely been limited to administration of RNA molecules. Hence, such administration must be performed repeatedly to have any sustained effect. The present inventors have developed a delivery mechanism that results in specific silencing of targeted genes through expression of small interfering RNA (siRNA). The inventors have markedly diminished expression of exogenous and endogenous genes *in vitro* and *in vivo* in brain and liver, and further apply this novel strategy to a model system of a major class of

neurodegenerative disorders, the polyglutamine diseases, to show reduced polyglutamine aggregation in cells. This strategy is generally useful in reducing expression of target genes in order to model biological processes or to provide therapy for dominant human diseases.

Disclosed herein is a strategy that results in substantial silencing of targeted alleles *via* siRNA. Use of this strategy results in markedly diminished *in vitro* and *in vivo* expression of targeted alleles. This strategy is useful in reducing expression of targeted alleles in order to model biological processes or to provide therapy for human diseases. For example, this strategy can be applied to a major class of neurodegenerative disorders, the polyglutamine diseases, as is demonstrated by the reduction of polyglutamine aggregation in cells following application of the strategy. As used herein the term "substantial silencing" means that the mRNA of the targeted allele is inhibited and/or degraded by the presence of the introduced siRNA, such that expression of the targeted allele is reduced by about 10% to 100% as compared to the level of expression seen when the siRNA is not present. Generally, when an allele is substantially silenced, it will have at least 40%, 50%, 60%, to 70%, *e.g.*, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, to 79%, generally at least 80%, *e.g.*, 81%-84%, at least 85%, *e.g.*, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or even 100% reduction expression as compared to when the siRNA is not present. As used herein the term "substantially normal activity" means the level of expression of an allele when an siRNA has not been introduced to a cell.

Dominantly inherited diseases are ideal candidates for siRNA-based therapy. To explore the utility of siRNA in inherited human disorders, the present inventors employed cellular models to test whether mutant alleles responsible for these dominantly-inherited human disorders could be specifically targeted. First, three classes of dominantly inherited, untreatable neurodegenerative diseases were examined: polyglutamine (polyQ) neurodegeneration in MJD/SCA3, Huntington's disease and frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17). Machado-

Joseph disease is also known as Spinocerebellar Ataxia Type 3 (The HUGO official name is MJD). The gene involved is MJD1, which encodes for the protein ataxin-3 (also called Mjd1p). Huntington's disease is due to expansion of the CAG repeat motif in exon 1 of huntingtin. In 38% of patients a polymorphism exists in exon 58 of the huntingtin gene, allowing for allele specific targeting. Frontotemporal dementia (sometimes with parkinonism, and linked to chromosome 17, so sometimes called FTDP-17) is due to mutations in the MAPT1 gene that encodes the protein tau.

The polyQ neurodegenerative disorders include at least nine diseases caused by CAG repeat expansions that encode polyQ in the disease protein. PolyQ expansion confers a dominant toxic property on the mutant protein that is associated with aberrant accumulation of the disease protein in neurons (Zoghbi, 2000). In FTDP-17, Tau mutations lead to the formation of neurofibrillary tangles accompanied by neuronal dysfunction and degeneration (Poorkaj, 1998; Hutton, 1998). The precise mechanisms by which these mutant proteins cause neuronal injury are unknown, but considerable evidence suggests that the abnormal proteins themselves initiate the pathogenic process (Zoghbi, 2000). Accordingly, eliminating expression of the mutant protein by siRNA or other means slows or prevents disease (Yamamoto, 2000). However, because many dominant disease genes also encode essential proteins (*e.g.* Nasir, 1995) siRNA-mediated approaches were developed that selectively inactivate mutant alleles, while allowing continued expression of the wild type proteins ataxin-3 and huntingtin.

Second, the dominantly-inherited disorder DYT1 dystonia was studied. DYT1 dystonia is also known as Torsion dystonia type 1, and is caused by a GAG deletion in the TOR1A gene encoding torsinA. DYT1 dystonia is the most common cause of primary generalized dystonia. DYT1 usually presents in childhood as focal dystonia that progresses to severe generalized disease (Fahn, 1998; Klein, 2002a). With one possible exception (Leung, 2001; Doheny, 2002; Klein, 2002), all cases of DYT1 result from a common GAG deletion in *TOR1A*, eliminating one of two adjacent glutamic acids near the C-terminus of the

protein TorsinA (TA) (Ozelius, 1997). Although the precise cellular function of TA is unknown, it seems clear that mutant TA (TAmut) acts through a dominant-negative or dominant-toxic mechanism (Breakefield, 2001).

Several characteristics of DYT1 make it an ideal disease in which to use siRNA-mediated gene silencing as therapy. Of greatest importance, the dominant nature of the disease suggests that a reduction in mutant TA, whatever the precise pathogenic mechanism proves to be, is helpful. Moreover, the existence of a single common mutation that deletes a full three nucleotides suggested it might be feasible to design siRNA that specifically targets the mutant allele and is applicable to all affected persons. Finally, there is no effective therapy for DYT1, a relentless and disabling disease.

As outlined in the strategy in Figure 11, the inventors developed siRNA that would specifically eliminate production of protein from the mutant allele. By exploiting the three base pair difference between wild type and mutant alleles, the inventors successfully silenced expression of the mutant protein (TAmut) without interfering with expression of the wild type protein (TAwt). Because TAwt may be an essential protein it is critically important that efforts be made to silence only the mutant allele. This allele-specific strategy has obvious therapeutic potential for DYT1 and represents a novel and powerful research tool with which to investigate the function of TA and its dysfunction in the disease state.

Expansions of poly-glutamine tracts in proteins that are expressed in the central nervous system can cause neurodegenerative diseases. Some neurodegenerative diseases are caused by a $(CAG)_n$ repeat that encodes poly-glutamine in a protein include Huntington disease (HD), spinocerebellar ataxia (SCA1, SCA2, SCA3, SCA6, SCA7), spinal and bulbar muscular atrophy (SBMA), and dentatorubropallidoluysian atrophy (DRPLA). In these diseases, the poly-glutamine expansion in a protein confers a novel toxic property upon the protein. Studies indicate that the toxic property is a tendency for the disease protein to misfold and form aggregates within neurons.

The gene involved in Huntington's disease (IT-15) is located at the end of the short arm of chromosome 4. This gene is designated HD and encodes the protein huntingtin (also known as Htt). A mutation occurs in the coding region of this gene and produces an unstable expanded trinucleotide repeat (cytosine-adenosine-guanosine), resulting in a protein with an expanded glutamate sequence. The normal and abnormal functions of this protein (termed huntingtin) are unknown. The abnormal huntingtin protein appears to accumulate in neuronal nuclei of transgenic mice, but the causal relationship of this accumulation to neuronal death is uncertain.

One of skill in the art can select additional target sites for generating siRNA specific for other alleles beyond those specifically described in the experimental examples. Such allele-specific siRNAs made be designed using the guidelines provided by Ambion (Austin, TX). Briefly, the target cDNA sequence is scanned for target sequences that had AA di-nucleotides. Sense and anti-sense oligonucleotides are generated to these targets (AA + 3' adjacent 19 nucleotides) that contained a G/C content of 35 to 55%. These sequences are then compared to others in the human genome database to minimize homology to other known coding sequences (BLAST search).

To accomplish intracellular expression of the therapeutic siRNA, an RNA molecule is constructed containing two complementary strands or a hairpin sequence (such as a 21-bp hairpin) representing sequences directed against the gene of interest. The siRNA, or a nucleic acid encoding the siRNA, is introduced to the target cell, such as a diseased brain cell. The siRNA reduces target mRNA and protein expression.

The construct encoding the therapeutic siRNA can be configured such that one or more strands of the siRNA are encoded by a nucleic acid that is immediately contiguous to a promoter. In one example, the promoter is a pol II promoter. If a pol II promoter is used in a particular construct, it is selected from readily available pol II promoters known in the art, depending on whether regulatable, inducible, tissue or cell-specific expression of the siRNA is desired.

The construct is introduced into the target cell, such as by injection, allowing for diminished target-gene expression in the cell.

It was surprising that a pol II promoter would be effective. While small RNAs with extensive secondary structure are routinely made from Pol III promoters, there is no *a priori* reason to assume that small interfering RNAs could be expressed from pol II promoters. Pol III promoters terminate in a short stretch of Ts (5 or 6), leaving a very small 3' end and allowing stabilization of secondary structure. Polymerase II transcription extends well past the coding and polyadenylation regions, after which the transcript is cleaved. Two adenylation steps occur, leaving a transcript with a tail of up to 200 As. This string of As would of course completely destabilize any small, 21 base pair hairpin. Therefore, in addition to modifying the promoter to minimize sequences between the transcription start site and the siRNA sequence (thereby stabilizing the hairpin), the inventors also extensively modified the polyadenylation sequence to test if a very short polyadenylation could occur. The results, which were not predicted from prior literature, showed that it could.

The present invention provides an expression cassette containing an isolated nucleic acid sequence encoding a small interfering RNA molecule (siRNA) targeted against a gene of interest. The siRNA may form a hairpin structure that contains a duplex structure and a loop structure. The loop structure may contain from 4 to 10 nucleotides, such as 4, 5 or 6 nucleotides. The duplex is less than 30 nucleotides in length, such as from 19 to 25 nucleotides. The siRNA may further contain an overhang region. Such an overhang may be a 3' overhang region or a 5' overhang region. The overhang region may be, for example, from 1 to 6 nucleotides in length. The expression cassette may further contain a pol II promoter, as described herein. Examples of pol II promoters include regulatable promoters and constitutive promoters. For example, the promoter may be a CMV or RSV promoter. The expression cassette may further contain a polyadenylation signal, such as a synthetic minimal polyadenylation signal. The nucleic acid sequence may further contain a marker gene or stuffer sequences. The expression cassette may be contained in a viral vector. An

appropriate viral vector for use in the present invention may be an adenoviral, lentiviral, adeno-associated viral (AAV), poliovirus, herpes simplex virus (HSV) or murine Maloney-based viral vector. The gene of interest may be a gene associated with a condition amenable to siRNA therapy. Examples of such conditions include neurodegenerative diseases, such as a trinucleotide-repeat disease (*e.g.*, polyglutamine repeat disease). Examples of these diseases include Huntington's disease or several spinocerebellar ataxias. Alternatively, the gene of interest may encode a ligand for a chemokine involved in the migration of a cancer cell, or a chemokine receptor.

The present invention also provides an expression cassette containing an isolated nucleic acid sequence encoding a first segment, a second segment located immediately 3' of the first segment, and a third segment located immediately 3' of the second segment, wherein the first and third segments are each less than 30 base pairs in length and each more than 10 base pairs in length, and wherein the sequence of the third segment is the complement of the sequence of the first segment, and wherein the isolated nucleic acid sequence functions as a small interfering RNA molecule (siRNA) targeted against a gene of interest. The expression cassette may be contained in a vector, such as a viral vector.

The present invention provides a method of reducing the expression of a gene product in a cell by contacting a cell with an expression cassette described above. It also provides a method of treating a patient by administering to the patient a composition of the expression cassette described above.

The present invention further provides a method of reducing the expression of a gene product in a cell by contacting a cell with an expression cassette containing an isolated nucleic acid sequence encoding a first segment, a second segment located immediately 3' of the first segment, and a third segment located immediately 3' of the second segment, wherein the first and third segments are each less than 30 base pairs in length and each more than 10 base pairs in length, and wherein the sequence of the third segment is the complement of the sequence of the first segment, and wherein the isolated nucleic acid

sequence functions as a small interfering RNA molecule (siRNA) targeted against a gene of interest.

The present method also provides a method of treating a patient, by administering to the patient a composition containing an expression cassette, wherein the expression cassette contains an isolated nucleic acid sequence encoding a first segment, a second segment located immediately 3' of the first segment, and a third segment located immediately 3' of the second segment, wherein the first and third segments are each less than 30 bases in length and each more than 10 bases in length, and wherein the sequence of the third segment is the complement of the sequence of the first segment, and wherein the isolated nucleic acid sequence functions as a small interfering RNA molecule (siRNA) targeted against a gene of interest.

I. Interfering RNA

A "small interfering RNA" or "short interfering RNA" or "siRNA" or "short hairpin RNA" or "shRNA" is a RNA duplex of nucleotides that is targeted to a nucleic acid sequence of interest, for example, a Huntington's Disease gene (also referred to as huntingtin, htt, or HD). As used herein, the term "siRNA" is a generic term that encompasses the subset of shRNAs. A "RNA duplex" refers to the structure formed by the complementary pairing between two regions of a RNA molecule. siRNA is "targeted" to a gene in that the nucleotide sequence of the duplex portion of the siRNA is complementary to a nucleotide sequence of the targeted gene. In certain embodiments, the siRNAs are targeted to the sequence encoding huntingtin. In some embodiments, the length of the duplex of siRNAs is less than 30 base pairs. In some embodiments, the duplex can be 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11 or 10 base pairs in length. In some embodiments, the length of the duplex is 19 to 25 base pairs in length. In certain embodiment, the length of the duplex is 19 or 21 base pairs in length. The RNA duplex portion of the siRNA can be part of a hairpin structure. In addition to the duplex portion, the hairpin structure may contain a loop portion positioned between the two sequences that form the duplex. The

loop can vary in length. In some embodiments the loop is 5, 6, 7, 8, 9, 10, 11, 12 or 13 nucleotides in length. In certain embodiments, the loop is 9 nucleotides in length. The hairpin structure can also contain 3′ or 5′ overhang portions. In some embodiments, the overhang is a 3′ or a 5′ overhang 0, 1, 2, 3, 4 or 5 nucleotides in length.

The siRNA can be encoded by a nucleic acid sequence, and the nucleic acid sequence can also include a promoter. The nucleic acid sequence can also include a polyadenylation signal. In some embodiments, the polyadenylation signal is a synthetic minimal polyadenylation signal.

"Knock-down," "knock-down technology" refers to a technique of gene silencing in which the expression of a target gene is reduced as compared to the gene expression prior to the introduction of the siRNA, which can lead to the inhibition of production of the target gene product. The term "reduced" is used herein to indicate that the target gene expression is lowered by 1-100%. In other words, the amount of RNA available for translation into a polypeptide or protein is minimized. For example, the amount of protein may be reduced by 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, or 99%. In some embodiments, the expression is reduced by about 90% (*i.e.*, only about 10% of the amount of protein is observed a cell as compared to a cell where siRNA molecules have not been administered). Knock-down of gene expression can be directed by the use of dsRNAs or siRNAs.

"RNA interference (RNAi)" is the process of sequence-specific, post-transcriptional gene silencing initiated by siRNA. During RNAi, siRNA induces degradation of target mRNA with consequent sequence-specific inhibition of gene expression. RNAi involving the use of siRNA has been successfully applied to knockdown the expression of specific genes in plants, D. melanogaster, C. elegans, trypanosomes, planaria, hydra, and several vertebrate species including the mouse. For a review of the mechanisms proposed to mediate RNAi, please refer to Bass *et al.*, 2001 Elbashir, 2001a, 2001b, 2001c; or Brantl, 2002.

According to a method of the present invention, the expression of huntingtin can be modified via RNAi. For example, the accumulation of huntingtin can be suppressed in a cell. The term "suppressing" refers to the diminution, reduction or elimination in the number or amount of transcripts present in a particular cell. For example, the accumulation of mRNA encoding huntingtin can be suppressed in a cell by RNA interference (RNAi), *e.g.*, the gene is silenced by sequence-specific double-stranded RNA (dsRNA), which is also called short interfering RNA (siRNA). These siRNAs can be two separate RNA molecules that have hybridized together, or they may be a single hairpin wherein two portions of a RNA molecule have hybridized together to form a duplex.

A mutant protein refers to the protein encoded by a gene having a mutation, *e.g.*, a missense or nonsense mutation in one or both alleles of huntingtin. A mutant huntingtin may be disease-causing, *i.e.*, may lead to a disease associated with the presence of huntingtin in an animal having either one or two mutant allele(s). The term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form, composed of monomers (nucleotides) containing a sugar, phosphate and a base that is either a purine or pyrimidine. Unless specifically limited, the term encompasses nucleic acids containing known analogs of natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also encompasses conservatively modified variants thereof (*e.g.*, degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, (1991); Ohtsuka *et al.*, (1985); Rossolini *et al.*, (1994)).

A "nucleic acid fragment" is a portion of a given nucleic acid molecule. Deoxyribonucleic acid (DNA) in the majority of organisms is the genetic

material while ribonucleic acid (RNA) is involved in the transfer of information contained within DNA into proteins.

The term "nucleotide sequence" refers to a polymer of DNA or RNA which can be single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases capable of incorporation into DNA or RNA polymers.

The terms "nucleic acid", "nucleic acid molecule", "nucleic acid fragment", "nucleic acid sequence or segment", or "polynucleotide" are used interchangeably and may also be used interchangeably with gene, cDNA, DNA and RNA encoded by a gene.

The invention encompasses isolated or substantially purified nucleic acid or protein compositions. In the context of the present invention, an "isolated" or "purified" DNA molecule or RNA molecule or an "isolated" or "purified" polypeptide is a DNA molecule, RNA molecule, or polypeptide that exists apart from its native environment and is therefore not a product of nature. An isolated DNA molecule, RNA molecule or polypeptide may exist in a purified form or may exist in a non-native environment such as, for example, a transgenic host cell. For example, an "isolated" or "purified" nucleic acid molecule or protein, or biologically active portion thereof, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. In one embodiment, an "isolated" nucleic acid is free of sequences that naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences that naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. A protein that is substantially free of cellular material includes preparations of protein or polypeptide having less than about 30%, 20%, 10%, or 5% (by dry weight) of contaminating protein. When the protein of the invention, or biologically active

portion thereof, is recombinantly produced, culture medium represents less than about 30%, 20%, 10%, or 5% (by dry weight) of chemical precursors or non-protein-of-interest chemicals. Fragments and variants of the disclosed nucleotide sequences and proteins or partial-length proteins encoded thereby are also encompassed by the present invention. By "fragment" or "portion" is meant a full length or less than full length of the nucleotide sequence encoding, or the amino acid sequence of, a polypeptide or protein.

The term "gene" is used broadly to refer to any segment of nucleic acid associated with a biological function. Thus, genes include coding sequences and/or the regulatory sequences required for their expression. For example, "gene" refers to a nucleic acid fragment that expresses mRNA, functional RNA, or specific protein, including regulatory sequences. "Genes" also include nonexpressed DNA segments that, for example, form recognition sequences for other proteins. "Genes" can be obtained from a variety of sources, including cloning from a source of interest or synthesizing from known or predicted sequence information, and may include sequences designed to have desired parameters. An "allele" is one of several alternative forms of a gene occupying a given locus on a chromosome.

"Naturally occurring," "native" or "wildtype" are used to describe an object that can be found in nature as distinct from being artificially produced. For example, a protein or nucleotide sequence present in an organism (including a virus), which can be isolated from a source in nature and which has not been intentionally modified by a person in the laboratory, is naturally occurring.

The term "chimeric" refers to a gene or DNA that contains 1) DNA sequences, including regulatory and coding sequences that are not found together in nature or 2) sequences encoding parts of proteins not naturally adjoined, or 3) parts of promoters that are not naturally adjoined. Accordingly, a chimeric gene may include regulatory sequences and coding sequences that are derived from different sources, or include regulatory sequences and coding sequences derived from the same source, but arranged in a manner different from that found in nature.

A "transgene" refers to a gene that has been introduced into the genome by transformation. Transgenes include, for example, DNA that is either heterologous or homologous to the DNA of a particular cell to be transformed. Additionally, transgenes may include native genes inserted into a non-native organism, or chimeric genes.

The term "endogenous gene" refers to a native gene in its natural location in the genome of an organism.

A "foreign" gene refers to a gene not normally found in the host organism that has been introduced by gene transfer.

The terms "protein," "peptide" and "polypeptide" are used interchangeably herein.

A "variant" of a molecule is a sequence that is substantially similar to the sequence of the native molecule. For nucleotide sequences, variants include those sequences that, because of the degeneracy of the genetic code, encode the identical amino acid sequence of the native protein. Naturally occurring allelic variants such as these can be identified with the use of molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques. Variant nucleotide sequences also include synthetically derived nucleotide sequences, such as those generated, for example, by using site-directed mutagenesis, which encode the native protein, as well as those that encode a polypeptide having amino acid substitutions. Generally, nucleotide sequence variants of the invention will have at least 40%, 50%, 60%, to 70%, *e.g.*, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, to 79%, generally at least 80%, *e.g.*, 81%-84%, at least 85%, *e.g.*, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, to 98%, sequence identity to the native (endogenous) nucleotide sequence.

"Conservatively modified variations" of a particular nucleic acid sequence refers to those nucleic acid sequences that encode identical or essentially identical amino acid sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given polypeptide. For instance, the codons CGT, CGC, CGA, CGG, AGA and

AGG all encode the amino acid arginine. Thus, at every position where an arginine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded protein. Such nucleic acid variations are "silent variations," which are one species of "conservatively modified variations." Every nucleic acid sequence described herein that encodes a polypeptide also describes every possible silent variation, except where otherwise noted. One of skill in the art will recognize that each codon in a nucleic acid (except ATG, which is ordinarily the only codon for methionine) can be modified to yield a functionally identical molecule by standard techniques. Accordingly, each "silent variation" of a nucleic acid that encodes a polypeptide is implicit in each described sequence.

"Recombinant DNA molecule" is a combination of DNA sequences that are joined together using recombinant DNA technology and procedures used to join together DNA sequences as described, for example, in Sambrook and Russell (2001).

The terms "heterologous gene", "heterologous DNA sequence", "exogenous DNA sequence", "heterologous RNA sequence", "exogenous RNA sequence" or "heterologous nucleic acid" each refer to a sequence that either originates from a source foreign to the particular host cell, or is from the same source but is modified from its original or native form. Thus, a heterologous gene in a host cell includes a gene that is endogenous to the particular host cell but has been modified through, for example, the use of DNA shuffling. The terms also include non-naturally occurring multiple copies of a naturally occurring DNA or RNA sequence. Thus, the terms refer to a DNA or RNA segment that is foreign or heterologous to the cell, or homologous to the cell but in a position within the host cell nucleic acid in which the element is not ordinarily found. Exogenous DNA segments are expressed to yield exogenous polypeptides.

A "homologous" DNA or RNA sequence is a sequence that is naturally associated with a host cell into which it is introduced.

"Wild-type" refers to the normal gene or organism found in nature.

"Genome" refers to the complete genetic material of an organism.

A "vector" is defined to include, *inter alia*, any viral vector, as well as any plasmid, cosmid, phage or binary vector in double or single stranded linear or circular form that may or may not be self transmissible or mobilizable, and that can transform prokaryotic or eukaryotic host either by integration into the cellular genome or exist extrachromosomally (*e.g.*, autonomous replicating plasmid with an origin of replication).

"Expression cassette" as used herein means a nucleic acid sequence capable of directing expression of a particular nucleotide sequence in an appropriate host cell, which may include a promoter operably linked to the nucleotide sequence of interest that may be operably linked to termination signals. It also may include sequences required for proper translation of the nucleotide sequence. The coding region usually codes for a protein of interest but may also code for a functional RNA of interest, for example an antisense RNA, a nontranslated RNA in the sense or antisense direction, or a siRNA. The expression cassette including the nucleotide sequence of interest may be chimeric. The expression cassette may also be one that is naturally occurring but has been obtained in a recombinant form useful for heterologous expression. The expression of the nucleotide sequence in the expression cassette may be under the control of a constitutive promoter or of an regulatable promoter that initiates transcription only when the host cell is exposed to some particular stimulus. In the case of a multicellular organism, the promoter can also be specific to a particular tissue or organ or stage of development.

Such expression cassettes can include a transcriptional initiation region linked to a nucleotide sequence of interest. Such an expression cassette is provided with a plurality of restriction sites for insertion of the gene of interest to be under the transcriptional regulation of the regulatory regions. The expression cassette may additionally contain selectable marker genes.

"Coding sequence" refers to a DNA or RNA sequence that codes for a specific amino acid sequence. It may constitute an "uninterrupted coding sequence", *i.e.*, lacking an intron, such as in a cDNA, or it may include one or

more introns bounded by appropriate splice junctions. An "intron" is a sequence of RNA that is contained in the primary transcript but is removed through cleavage and re-ligation of the RNA within the cell to create the mature mRNA that can be translated into a protein.

The term "open reading frame" (ORF) refers to the sequence between translation initiation and termination codons of a coding sequence. The terms "initiation codon" and "termination codon" refer to a unit of three adjacent nucleotides (a 'codon') in a coding sequence that specifies initiation and chain termination, respectively, of protein synthesis (mRNA translation).

"Functional RNA" refers to sense RNA, antisense RNA, ribozyme RNA, siRNA, or other RNA that may not be translated but yet has an effect on at least one cellular process.

The term "RNA transcript" refers to the product resulting from RNA polymerase catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA" (mRNA) refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a single- or a double-stranded DNA that is complementary to and derived from mRNA.

"Regulatory sequences" and "suitable regulatory sequences" each refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences include enhancers, promoters, translation leader sequences, introns, and polyadenylation signal sequences. They include natural and synthetic sequences as well as sequences that may be a combination of synthetic and natural sequences. As is noted above, the term "suitable regulatory sequences" is not limited to promoters. However, some suitable regulatory sequences useful in the present invention will

include, but are not limited to constitutive promoters, tissue-specific promoters, development-specific promoters, regulatable promoters and viral promoters. Examples of promoters that may be used in the present invention include CMV, RSV, pol II and pol III promoters.

"5' non-coding sequence" refers to a nucleotide sequence located 5' (upstream) to the coding sequence. It is present in the fully processed mRNA upstream of the initiation codon and may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency (Turner *et al.*, 1995).

"3' non-coding sequence" refers to nucleotide sequences located 3' (downstream) to a coding sequence and may include polyadenylation signal sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor.

The term "translation leader sequence" refers to that DNA sequence portion of a gene between the promoter and coding sequence that is transcribed into RNA and is present in the fully processed mRNA upstream (5') of the translation start codon. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency.

The term "mature" protein refers to a post-translationally processed polypeptide without its signal peptide. "Precursor" protein refers to the primary product of translation of an mRNA. "Signal peptide" refers to the amino terminal extension of a polypeptide, which is translated in conjunction with the polypeptide forming a precursor peptide and which is required for its entrance into the secretory pathway. The term "signal sequence" refers to a nucleotide sequence that encodes the signal peptide.

"Promoter" refers to a nucleotide sequence, usually upstream (5') to its coding sequence, which directs and/or controls the expression of the coding sequence by providing the recognition for RNA polymerase and other factors required for proper transcription. "Promoter" includes a minimal promoter that is

a short DNA sequence comprised of a TATA-box and other sequences that serve to specify the site of transcription initiation, to which regulatory elements are added for control of expression. "Promoter" also refers to a nucleotide sequence that includes a minimal promoter plus regulatory elements that is capable of controlling the expression of a coding sequence or functional RNA. This type of promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a DNA sequence that can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue specificity of a promoter. It is capable of operating in both orientations (normal or flipped), and is capable of functioning even when moved either upstream or downstream from the promoter. Both enhancers and other upstream promoter elements bind sequence-specific DNA-binding proteins that mediate their effects. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even be comprised of synthetic DNA segments. A promoter may also contain DNA sequences that are involved in the binding of protein factors that control the effectiveness of transcription initiation in response to physiological or developmental conditions.

The "initiation site" is the position surrounding the first nucleotide that is part of the transcribed sequence, which is also defined as position +1. With respect to this site all other sequences of the gene and its controlling regions are numbered. Downstream sequences (*i.e.*, further protein encoding sequences in the 3' direction) are denominated positive, while upstream sequences (mostly of the controlling regions in the 5' direction) are denominated negative.

Promoter elements, particularly a TATA element, that are inactive or that have greatly reduced promoter activity in the absence of upstream activation are referred to as "minimal or core promoters." In the presence of a suitable transcription factor, the minimal promoter functions to permit transcription. A "minimal or core promoter" thus consists only of all basal elements needed for transcription initiation, *e.g.*, a TATA box and/or an initiator.

"Constitutive expression" refers to expression using a constitutive or regulated promoter. "Conditional" and "regulated expression" refer to expression controlled by a regulated promoter.

"Operably-linked" refers to the association of nucleic acid sequences on single nucleic acid fragment so that the function of one of the sequences is affected by another. For example, a regulatory DNA sequence is said to be "operably linked to" or "associated with" a DNA sequence that codes for an RNA or a polypeptide if the two sequences are situated such that the regulatory DNA sequence affects expression of the coding DNA sequence (*i.e.*, that the coding sequence or functional RNA is under the transcriptional control of the promoter). Coding sequences can be operably-linked to regulatory sequences in sense or antisense orientation.

"Expression" refers to the transcription and/or translation of an endogenous gene, heterologous gene or nucleic acid segment, or a transgene in cells. For example, in the case of siRNA constructs, expression may refer to the transcription of the siRNA only. In addition, expression refers to the transcription and stable accumulation of sense (mRNA) or functional RNA. Expression may also refer to the production of protein.

"Altered levels" refers to the level of expression in transgenic cells or organisms that differs from that of normal or untransformed cells or organisms.

"Overexpression" refers to the level of expression in transgenic cells or organisms that exceeds levels of expression in normal or untransformed cells or organisms.

"Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of protein from an endogenous gene or a transgene.

"Transcription stop fragment" refers to nucleotide sequences that contain one or more regulatory signals, such as polyadenylation signal sequences, capable of terminating transcription. Examples include the 3' non-regulatory regions of genes encoding nopaline synthase and the small subunit of ribulose bisphosphate carboxylase.

"Translation stop fragment" refers to nucleotide sequences that contain one or more regulatory signals, such as one or more termination codons in all three frames, capable of terminating translation. Insertion of a translation stop fragment adjacent to or near the initiation codon at the 5' end of the coding sequence will result in no translation or improper translation. Excision of the translation stop fragment by site-specific recombination will leave a site-specific sequence in the coding sequence that does not interfere with proper translation using the initiation codon.

The terms "*cis*-acting sequence" and "*cis*-acting element" refer to DNA or RNA sequences whose functions require them to be on the same molecule. An example of a *cis*-acting sequence on the replicon is the viral replication origin.

The terms "*trans*-acting sequence" and "*trans*-acting element" refer to DNA or RNA sequences whose function does not require them to be on the same molecule.

"Chromosomally-integrated" refers to the integration of a foreign gene or nucleic acid construct into the host DNA by covalent bonds. Where genes are not "chromosomally integrated" they may be "transiently expressed." Transient expression of a gene refers to the expression of a gene that is not integrated into the host chromosome but functions independently, either as part of an autonomously replicating plasmid or expression cassette, for example, or as part of another biological system such as a virus.

The following terms are used to describe the sequence relationships between two or more nucleic acids or polynucleotides: (a) "reference sequence", (b) "comparison window", (c) "sequence identity", (d) "percentage of sequence identity", and (e) "substantial identity".

(a) As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence.

(b) As used herein, "comparison window" makes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Generally, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of sequences for comparison are well-known in the art. Thus, the determination of percent identity between any two sequences can be accomplished using a mathematical algorithm. Non-limiting examples of such mathematical algorithms are the algorithm of Myers and Miller (1988); the local homology algorithm of Smith *et al.* (1981); the homology alignment algorithm of Needleman and Wunsch (1970); the search-for-similarity-method of Pearson and Lipman (1988); the algorithm of Karlin and Altschul (1990), modified as in Karlin and Altschul (1993).

Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, but are not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Version 8 (available from Genetics Computer Group (GCG), 575 Science Drive, Madison, Wisconsin, USA). Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins *et al.* (1988); Higgins *et al.* (1989); Corpet *et al.* (1988); Huang *et al.* (1992); and Pearson *et al.* (1994). The ALIGN program is based on the algorithm of Myers and Miller, *supra*. The BLAST programs of Altschul *et al.* (1990), are based on the algorithm of Karlin and Altschul *supra*.

Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold. These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when the cumulative alignment score falls off by the quantity X from its maximum achieved value, the cumulative score goes to zero or below due to the accumulation of one or more negative-scoring residue alignments, or the end of either sequence is reached.

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a test nucleic acid sequence is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid sequence to the reference nucleic acid sequence is less than about 0.1, less than about 0.01, or even less than about 0.001.

To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul *et al.* (1997). Alternatively, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. See Altschul *et al.*,

supra. When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs (*e.g.* BLASTN for nucleotide sequences, BLASTX for proteins) can be used. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix. See http://www.ncbi.n1m.nih.gov. Alignment may also be performed manually by inspection.

For purposes of the present invention, comparison of nucleotide sequences for determination of percent sequence identity to the promoter sequences disclosed herein is made using the BlastN program (version 1.4.7 or later) with its default parameters or any equivalent program. By "equivalent program" is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by the preferred program.

(c) As used herein, "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences makes reference to a specified percentage of residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window, as measured by sequence comparison algorithms or by visual inspection. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (*e.g.*, charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity." Means for making this adjustment are well known to

those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, *e.g.*, as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

(d) As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

(e)(i) The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, or 79%, or at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, or 89%, or at least 90%, 91%, 92%, 93%, or 94%, or even at least 95%, 96%, 97%, 98%, or 99% sequence identity, compared to a reference sequence using one of the alignment programs described using standard parameters. One of skill in the art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning, and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 70%, at least 80%, 90%, or even at least 95%.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. However, stringent conditions encompass temperatures in the range of about 1°C to about 20°C, depending upon the desired degree of stringency as otherwise qualified herein. Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides they encode are substantially identical. This may occur, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. One indication that two nucleic acid sequences are substantially identical is when the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid.

(e)(ii) The term "substantial identity" in the context of a peptide indicates that a peptide comprises a sequence with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, or 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, or 89%, or at least 90%, 91%, 92%, 93%, or 94%, or even, 95%, 96%, 97%, 98% or 99%, sequence identity to the reference sequence over a specified comparison window. Optimal alignment may be conducted using the homology alignment algorithm of Needleman and Wunsch (1970). An indication that two peptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a peptide is substantially identical to a second peptide, for example, where the two peptides differ only by a conservative substitution.

For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

As noted above, another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions. The phrase "hybridizing specifically to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (*e.g.*, total cellular) DNA or RNA. "Bind(s) substantially" refers to complementary hybridization between a probe nucleic acid and a target nucleic acid and embraces minor mismatches that can be accommodated by reducing the stringency of the hybridization media to achieve the desired detection of the target nucleic acid sequence.

"Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and Northern hybridizations are sequence dependent, and are different under different environmental parameters. Longer sequences hybridize specifically at higher temperatures. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl (1984); T_m 81.5°C + 16.6 (log M) +0.41 (%GC) - 0.61 (% form) - 500/L; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. T_m is reduced by about 1°C for each 1% of mismatching; thus, T_m, hybridization, and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with >90% identity are sought, the T_m can be decreased 10°C. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than

the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point (T_m). Using the equation, hybridization and wash compositions, and desired T, those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T of less than 45°C (aqueous solution) or 32°C (formamide solution), the SSC concentration may be increased so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993). Generally, highly stringent hybridization and wash conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH.

An example of highly stringent wash conditions is 0.15 M NaCl at 72°C for about 15 minutes. An example of stringent wash conditions is a 0.2X SSC wash at 65°C for 15 minutes (see, Sambrook and Russell, *infra*, for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example medium stringency wash for a duplex of, *e.g.*, more than 100 nucleotides, is 1X SSC at 45°C for 15 minutes. An example low stringency wash for a duplex of, *e.g.*, more than 100 nucleotides, is 4-6X SSC at 40°C for 15 minutes. For short probes (*e.g.*, about 10 to 50 nucleotides), stringent conditions typically involve salt concentrations of less than about 1.5 M, about 0.01 to 1.0 M, Na ion concentration (or other salts) at pH 7.0 to 8.3, and the temperature is typically at least about 30°C and at least about 60°C for long probes (*e.g.*, >50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. In general, a signal to noise ratio of 2X (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the proteins that they encode are substantially identical. This occurs, *e.g.*, when a copy of a

nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

Very stringent conditions are selected to be equal to the T_m for a particular probe. An example of stringent conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or Northern blot is 50% formamide, *e.g.*, hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1M NaCl, 1% SDS (sodium dodecyl sulfate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1.0 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C.

By "variant" polypeptide is intended a polypeptide derived from the native protein by deletion (also called "truncation") or addition of one or more amino acids to the N-terminal and/or C-terminal end of the native protein; deletion or addition of one or more amino acids at one or more sites in the native protein; or substitution of one or more amino acids at one or more sites in the native protein. Such variants may result from, for example, genetic polymorphism or from human manipulation. Methods for such manipulations are generally known in the art.

Thus, the polypeptides of the invention may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants of the polypeptides can be prepared by mutations in the DNA. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Kunkel (1985); Kunkel *et al.* (1987); U. S. Patent No. 4,873,192; Walker and Gaastra (1983), and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of

Dayhoff *et al.* (1978). Conservative substitutions, such as exchanging one amino acid with another having similar properties, may be used.

Thus, the genes and nucleotide sequences of the invention include both the naturally occurring sequences as well as variant forms. Likewise, the polypeptides of the invention encompass both naturally occurring proteins as well as variations and modified forms thereof. Such variants will continue to possess the desired activity. The deletions, insertions, and substitutions of the polypeptide sequence encompassed herein are not expected to produce radical changes in the characteristics of the polypeptide. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays.

Individual substitutions deletions or additions that alter, add or delete a single amino acid or a small percentage of amino acids (typically less than 5%, more typically less than 1%) in an encoded sequence are "conservatively modified variations," where the alterations result in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. The following five groups each contain amino acids that are conservative substitutions for one another: Aliphatic: Glycine (G), Alanine (A), Valine (V), Leucine (L), Isoleucine (I); Aromatic: Phenylalanine (F), Tyrosine (Y), Tryptophan (W); Sulfur-containing: Methionine (M), Cysteine (C); Basic: Arginine (R), Lysine (K), Histidine (H); Acidic: Aspartic acid (D), Glutamic acid (E), Asparagine (N), Glutamine (Q). In addition, individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids in an encoded sequence are also "conservatively modified variations."

The term "transformation" refers to the transfer of a nucleic acid fragment into the genome of a host cell, resulting in genetically stable inheritance. A "host cell" is a cell that has been transformed, or is capable of transformation, by an exogenous nucleic acid molecule. Host cells containing the

transformed nucleic acid fragments are referred to as "transgenic" cells, and organisms comprising transgenic cells are referred to as "transgenic organisms".

"Transformed", "transduced", "transgenic", and "recombinant" refer to a host cell or organism into which a heterologous nucleic acid molecule has been introduced. The nucleic acid molecule can be stably integrated into the genome generally known in the art and are disclosed in Sambrook and Russell, *infra*. See also Innis *et al.* (1995); and Gelfand (1995); and Innis and Gelfand (1999). Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially mismatched primers, and the like. For example, "transformed," "transformant," and "transgenic" cells have been through the transformation process and contain a foreign gene integrated into their chromosome. The term "untransformed" refers to normal cells that have not been through the transformation process.

A "transgenic" organism is an organism having one or more cells that contain an expression vector.

"Genetically altered cells" denotes cells which have been modified by the introduction of recombinant or heterologous nucleic acids (*e.g.*, one or more DNA constructs or their RNA counterparts) and further includes the progeny of such cells which retain part or all of such genetic modification.

The term "fusion protein" is intended to describe at least two polypeptides, typically from different sources, which are operably linked. With regard to polypeptides, the term operably linked is intended to mean that the two polypeptides are connected in a manner such that each polypeptide can serve its intended function. Typically, the two polypeptides are covalently attached through peptide bonds. The fusion protein is produced by standard recombinant DNA techniques. For example, a DNA molecule encoding the first polypeptide is ligated to another DNA molecule encoding the second polypeptide, and the resultant hybrid DNA molecule is expressed in a host cell to produce the fusion protein. The DNA molecules are ligated to each other in a 5' to 3' orientation

such that, after ligation, the translational frame of the encoded polypeptides is not altered (*i.e.*, the DNA molecules are ligated to each other in-frame).

As used herein, the term "derived" or "directed to" with respect to a nucleotide molecule means that the molecule has complementary sequence identity to a particular molecule of interest.

"Gene silencing" refers to the suppression of gene expression, *e.g.*, transgene, heterologous gene and/or endogenous gene expression. Gene silencing may be mediated through processes that affect transcription and/or through processes that affect post-transcriptional mechanisms. In some embodiments, gene silencing occurs when siRNA initiates the degradation of the mRNA of a gene of interest in a sequence-specific manner *via* RNA interference (for a review, see Brantl, 2002). In some embodiments, gene silencing may be allele-specific. "Allele-specific" gene silencing refers to the specific silencing of one allele of a gene.

"Knock-down," "knock-down technology" refers to a technique of gene silencing in which the expression of a target gene is reduced as compared to the gene expression prior to the introduction of the siRNA, which can lead to the inhibition of production of the target gene product. The term "reduced" is used herein to indicate that the target gene expression is lowered by 1-100%. For example, the expression may be reduced by 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, or even 99%. Knock-down of gene expression can be directed by the use of dsRNAs or siRNAs. For example, "RNA interference (RNAi)," which can involve the use of siRNA, has been successfully applied to knockdown the expression of specific genes in plants, *D. melanogaster, C. elegans*, trypanosomes, planaria, hydra, and several vertebrate species including the mouse. For a review of the mechanisms proposed to mediate RNAi, please refer to Bass *et al.*, 2001, Elbashir *et al.*, 2001 or Brantl 2002.

"RNA interference (RNAi)" is the process of sequence-specific, post-transcriptional gene silencing initiated by siRNA. RNAi is seen in a number of organisms such as *Drosophila*, nematodes, fungi and plants, and is believed to be involved in anti-viral defense, modulation of transposon activity, and

regulation of gene expression. During RNAi, siRNA induces degradation of target mRNA with consequent sequence-specific inhibition of gene expression.

A "small interfering" or "short interfering RNA" or siRNA is a RNA duplex of nucleotides that is targeted to a gene interest. A "RNA duplex" refers to the structure formed by the complementary pairing between two regions of a RNA molecule. siRNA is "targeted" to a gene in that the nucleotide sequence of the duplex portion of the siRNA is complementary to a nucleotide sequence of the targeted gene. In some embodiments, the length of the duplex of siRNAs is less than 30 nucleotides. In some embodiments, the duplex can be 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11 or 10 nucleotides in length. In some embodiments, the length of the duplex is 19 - 25 nucleotides in length. The RNA duplex portion of the siRNA can be part of a hairpin structure. In addition to the duplex portion, the hairpin structure may contain a loop portion positioned between the two sequences that form the duplex. The loop can vary in length. In some embodiments the loop is 5, 6, 7, 8, 9, 10, 11, 12 or 13 nucleotides in length. The hairpin structure can also contain 3' or 5' overhang portions. In some embodiments, the overhang is a 3' or a 5' overhang 0, 1, 2, 3, 4 or 5 nucleotides in length. Examples of shRNA specific for huntingin are encoded by the DNA sequences provided in Figure 20. The "sense" and "antisense" sequences can be used with or without the loop region indicated to form siRNA molecules. Other loop regions can be substituted for the examples provided in this chart. As used herein, the term siRNA is meant to be equivalent to other terms used to describe nucleic acid molecules that are capable of mediating sequence specific RNAi, for example, double-stranded RNA (dsRNA), micro-RNA (miRNA), short hairpin RNA (shRNA), short interfering oligonucleotide, short interfering nucleic acid, post-transcriptional gene silencing RNA (ptgsRNA), and others. In addition, as used herein, the term RNAi is meant to be equivalent to other terms used to describe sequence specific RNA interference, such as post transcriptional gene silencing, translational inhibition, or epigenetic silencing. For example, siRNA molecules of the invention can be used to epigenetically silence genes at both the post-transcriptional level or the

pre-transcriptional level. In a non-limiting example, epigenetic modulation of gene expression by siRNA molecules of the invention can result from siRNA mediated modification of chromatin structure or methylation pattern to alter gene expression (see, for example, Verdel *et al.,* 2004, *Science,* 303, 672-676; Pal-Bhadra *et al.,* 2004, *Science,* 303, 669-672; Allshire, 2002, *Science,* 297, 1818-1819; Volpe *et al.,* 2002, *Science,* 297, 1833-1837; Jenuwein, 2002, *Science,* 297, 2215-2218; and Hall *et al.,* 2002, *Science,* 297, 2232-2237). In another non-limiting example, modulation of gene expression by siRNA molecules of the invention can result from siRNA mediated cleavage of RNA (either coding or non-coding RNA) via RISC, or alternately, translational inhibition as is known in the art.

The siRNA can be encoded by a nucleic acid sequence, and the nucleic acid sequence can also include a promoter. The nucleic acid sequence can also include a polyadenylation signal. In some embodiments, the polyadenylation signal is a synthetic minimal polyadenylation signal.

"Treating" as used herein refers to ameliorating at least one symptom of, curing and/or preventing the development of a disease or a condition.

"Neurological disease" and "neurological disorder" refer to both hereditary and sporadic conditions that are characterized by nervous system dysfunction, and which may be associated with atrophy of the affected central or peripheral nervous system structures, or loss of function without atrophy. A neurological disease or disorder that results in atrophy is commonly called a "neurodegenerative disease" or "neurodegenerative disorder." Neurodegenerative diseases and disorders include, but are not limited to, amyotrophic lateral sclerosis (ALS), hereditary spastic hemiplegia, primary lateral sclerosis, spinal muscular atrophy, Kennedy's disease, Alzheimer's disease, Parkinson's disease, multiple sclerosis, and repeat expansion neurodegenerative diseases, *e.g.*, diseases associated with expansions of trinucleotide repeats such as polyglutamine (polyQ) repeat diseases, *e.g.*, Huntington's disease (HD), spinocerebellar ataxia (SCA1, SCA2, SCA3, SCA6, SCA7, and SCA17), spinal and bulbar muscular atrophy (SBMA),

dentatorubropallidoluysian atrophy (DRPLA). An example of a neurological disorder that does not appear to result in atrophy is DYT1 dystonia.

The siRNAs of the present invention can be generated by any method known to the art, for example, by in vitro transcription, recombinantly, or by synthetic means. In one example, the siRNAs can be generated in vitro by using a recombinant enzyme, such as T7 RNA polymerase, and DNA oligonucleotide templates.

II. Nucleic Acid Molecules of the Invention

Sources of nucleotide sequences from which the present nucleic acid molecules can be obtained include any vertebrate, such as mammalian, cellular source.

As discussed above, the terms "isolated and/or purified" refer to *in vitro* isolation of a nucleic acid, *e.g.*, a DNA or RNA molecule from its natural cellular environment, and from association with other components of the cell, such as nucleic acid or polypeptide, so that it can be sequenced, replicated, and/or expressed. For example, "isolated nucleic acid" may be a DNA molecule containing less than 31 sequential nucleotides that is transcribed into an siRNA. Such an isolated siRNA may, for example, form a hairpin structure with a duplex 21 base pairs in length that is complementary or hybridizes to a sequence in a gene of interest, and remains stably bound under stringent conditions (as defined by methods well known in the art, *e.g.*, in Sambrook and Russell, 2001). Thus, the RNA or DNA is "isolated" in that it is free from at least one contaminating nucleic acid with which it is normally associated in the natural source of the RNA or DNA and is substantially free of any other mammalian RNA or DNA. The phrase "free from at least one contaminating source nucleic acid with which it is normally associated" includes the case where the nucleic acid is reintroduced into the source or natural cell but is in a different chromosomal location or is otherwise flanked by nucleic acid sequences not normally found in the source cell, *e.g.*, in a vector or plasmid.

In addition to a DNA sequence encoding a siRNA, the nucleic acid molecules of the invention include double-stranded interfering RNA molecules, which are also useful to inhibit expression of a target gene. In certain embodiment of the invention, siRNAs are employed to inhibit expression of a target gene. By "inhibit expression" is meant to reduce, diminish or suppress expression of a target gene. Expression of a target gene may be inhibited via "gene silencing." Gene silencing refers to the suppression of gene expression, *e.g.*, transgene, heterologous gene and/or endogenous gene expression, which may be mediated through processes that affect transcription and/or through processes that affect post-transcriptional mechanisms. In some embodiments, gene silencing occurs when siRNA initiates the degradation of the mRNA transcribed from a gene of interest in a sequence-specific manner via RNA interference, thereby preventing translation of the gene's product (for a review, see Brantl, 2002).

As used herein, the term "recombinant nucleic acid", *e.g.*, "recombinant DNA sequence or segment" refers to a nucleic acid, *e.g.*, to DNA, that has been derived or isolated from any appropriate cellular source, that may be subsequently chemically altered *in vitro*, so that its sequence is not naturally occurring, or corresponds to naturally occurring sequences that are not positioned as they would be positioned in a genome which has not been transformed with exogenous DNA. An example of preselected DNA "derived" from a source, would be a DNA sequence that is identified as a useful fragment within a given organism, and which is then chemically synthesized in essentially pure form. An example of such DNA "isolated" from a source would be a useful DNA sequence that is excised or removed from said source by chemical means, *e.g.*, by the use of restriction endonucleases, so that it can be further manipulated, *e.g.*, amplified, for use in the invention, by the methodology of genetic engineering.

Thus, recovery or isolation of a given fragment of DNA from a restriction digest can employ separation of the digest on polyacrylamide or agarose gel by electrophoresis, identification of the fragment of interest by

comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. See Lawn *et al.* (1981), and Goeddel *et al.* (1980). Therefore, "recombinant DNA" includes completely synthetic DNA sequences, semi-synthetic DNA sequences, DNA sequences isolated from biological sources, and DNA sequences derived from RNA, as well as mixtures thereof.

Nucleic acid molecules having base substitutions (*i.e.*, variants) are prepared by a variety of methods known in the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of an earlier prepared variant or a non-variant version of the nucleic acid molecule.

Oligonucleotide-mediated mutagenesis is a method for preparing substitution variants. This technique is known in the art as described by Adelman *et al.* (1983). Briefly, nucleic acid encoding a siRNA can be altered by hybridizing an oligonucleotide encoding the desired mutation to a DNA template, where the template is the single-stranded form of a plasmid or bacteriophage containing the unaltered or native gene sequence. After hybridization, a DNA polymerase is used to synthesize an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in the nucleic acid encoding siRNA. Generally, oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides that are completely complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridize properly to the single-stranded DNA template molecule. The oligonucleotides are readily synthesized using techniques known in the art such as that described by Crea *et al.* (1978).

The DNA template can be generated by those vectors that are either derived from bacteriophage M13 vectors (the commercially available M13mp18

and M13mp19 vectors are suitable), or those vectors that contain a single-stranded phage origin of replication as described by Viera *et al.* (1987). Thus, the DNA that is to be mutated may be inserted into one of these vectors to generate single-stranded template. Production of the single-stranded template is described in Chapter 3 of Sambrook and Russell, 2001. Alternatively, single-stranded DNA template may be generated by denaturing double-stranded plasmid (or other) DNA using standard techniques.

For alteration of the native DNA sequence (to generate amino acid sequence variants, for example), the oligonucleotide is hybridized to the single-stranded template under suitable hybridization conditions. A DNA polymerizing enzyme, usually the Klenow fragment of DNA polymerase I, is then added to synthesize the complementary strand of the template using the oligonucleotide as a primer for synthesis. A heteroduplex molecule is thus formed such that one strand of DNA encodes the mutated form of the DNA, and the other strand (the original template) encodes the native, unaltered sequence of the DNA. This heteroduplex molecule is then transformed into a suitable host cell, usually a prokaryote such as *E. coli* JM101. After the cells are grown, they are plated onto agarose plates and screened using the oligonucleotide primer radiolabeled with 32-phosphate to identify the bacterial colonies that contain the mutated DNA. The mutated region is then removed and placed in an appropriate vector, generally an expression vector of the type typically employed for transformation of an appropriate host.

The method described immediately above may be modified such that a homoduplex molecule is created wherein both strands of the plasmid contain the mutations(s). The modifications are as follows: The single-stranded oligonucleotide is annealed to the single-stranded template as described above. A mixture of three deoxyribonucleotides, deoxyriboadenosine (dATP), deoxyriboguanosine (dGTP), and deoxyribothymidine (dTTP), is combined with a modified thiodeoxyribocytosine called dCTP-(*S) (which can be obtained from the Amersham Corporation). This mixture is added to the template-oligonucleotide complex. Upon addition of DNA polymerase to this

mixture, a strand of DNA identical to the template except for the mutated bases is generated. In addition, this new strand of DNA will contain dCTP-(*S) instead of dCTP, which serves to protect it from restriction endonuclease digestion.

After the template strand of the double-stranded heteroduplex is nicked with an appropriate restriction enzyme, the template strand can be digested with ExoIII nuclease or another appropriate nuclease past the region that contains the site(s) to be mutagenized. The reaction is then stopped to leave a molecule that is only partially single-stranded. A complete double-stranded DNA homoduplex is then formed using DNA polymerase in the presence of all four deoxyribonucleotide triphosphates, ATP, and DNA ligase. This homoduplex molecule can then be transformed into a suitable host cell such as *E. coli* JM101.

III. Expression Cassettes of the Invention

To prepare expression cassettes, the recombinant DNA sequence or segment may be circular or linear, double-stranded or single-stranded. Generally, the DNA sequence or segment is in the form of chimeric DNA, such as plasmid DNA or a vector that can also contain coding regions flanked by control sequences that promote the expression of the recombinant DNA present in the resultant transformed cell.

A "chimeric" vector or expression cassette, as used herein, means a vector or cassette including nucleic acid sequences from at least two different species, or has a nucleic acid sequence from the same species that is linked or associated in a manner that does not occur in the "native" or wild type of the species.

Aside from recombinant DNA sequences that serve as transcription units for an RNA transcript , or portions thereof, a portion of the recombinant DNA may be untranscribed, serving a regulatory or a structural function. For example, the recombinant DNA may have a promoter that is active in mammalian cells.

Other elements functional in the host cells, such as introns, enhancers, polyadenylation sequences and the like, may also be a part of the recombinant

DNA. Such elements may or may not be necessary for the function of the DNA, but may provide improved expression of the DNA by affecting transcription, stability of the siRNA, or the like. Such elements may be included in the DNA as desired to obtain the optimal performance of the siRNA in the cell.

Control sequences are DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotic cells, for example, include a promoter, and optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Operably linked nucleic acids are nucleic acids placed in a functional relationship with another nucleic acid sequence. For example, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, operably linked DNA sequences are DNA sequences that are linked are contiguous. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accord with conventional practice.

The recombinant DNA to be introduced into the cells may contain either a selectable marker gene or a reporter gene or both to facilitate identification and selection of expressing cells from the population of cells sought to be transfected or infected through viral vectors. In other embodiments, the selectable marker may be carried on a separate piece of DNA and used in a co-transfection procedure. Both selectable markers and reporter genes may be flanked with appropriate regulatory sequences to enable expression in the host cells. Useful selectable markers are known in the art and include, for example, antibiotic-resistance genes, such as *neo* and the like.

Reporter genes are used for identifying potentially transfected cells and for evaluating the functionality of regulatory sequences. Reporter genes that encode for easily assayable proteins are well known in the art. In general, a

reporter gene is a gene that is not present in or expressed by the recipient organism or tissue and that encodes a protein whose expression is manifested by some easily detectable property, *e.g.*, enzymatic activity. For example, reporter genes include the chloramphenicol acetyl transferase gene (cat) from Tn9 of *E. coli* and the luciferase gene from firefly *Photinus pyralis*. Expression of the reporter gene is assayed at a suitable time after the DNA has been introduced into the recipient cells.

In order to prevent any packaging of AAV genomic sequences containing the rep and cap genes, a plasmid containing the rep and cap DNA fragment can be modified by the inclusion of a stuffer fragment into the AAV genome which causes the DNA to exceed the length for optimal packaging. Thus, in certain embodiments, the helper fragment is not packaged into AAV virions. This is a safety feature, ensuring that only a recombinant AAV vector genome that does not exceed optimal packaging size is packaged into virions. An AAV helper fragment that incorporates a stuffer sequence can exceed the wild-type genome length of 4.6 kb, and lengths above 105% of the wild-type will generally not be packaged. The stuffer fragment can be derived from, for example, such non-viral sources as the Lac-Z or beta-galactosidase gene.

The general methods for constructing recombinant DNA that can transfect target cells are well known to those skilled in the art, and the same compositions and methods of construction may be utilized to produce the DNA useful herein. For example, Sambrook and Russell, *infra*, provides suitable methods of construction.

The recombinant DNA can be readily introduced into the host cells, *e.g.*, mammalian, bacterial, yeast or insect cells by transfection with an expression vector composed of DNA encoding the siRNA by any procedure useful for the introduction into a particular cell, *e.g.*, physical or biological methods, to yield a cell having the recombinant DNA stably integrated into its genome or existing as a episomal element, so that the DNA molecules, or sequences of the present invention are expressed by the host cell. The DNA is introduced into host cells via a vector. The host cell is may be of eukaryotic origin, *e.g.*, plant,

mammalian, insect, yeast or fungal sources, but host cells of non-eukaryotic origin may also be employed.

Physical methods to introduce a preselected DNA into a host cell include calcium phosphate precipitation, lipofection, particle bombardment, microinjection, electroporation, and the like. Biological methods to introduce the DNA of interest into a host cell include the use of DNA and RNA viral vectors. For mammalian gene therapy, as described hereinbelow, it is desirable to use an efficient means of inserting a copy gene into the host genome. Viral vectors, and especially retroviral vectors, have become the most widely used method for inserting genes into mammalian, *e.g.*, human cells. Other viral vectors can be derived from poxviruses, herpes simplex virus I, adenoviruses and adeno-associated viruses, and the like. See, for example, U.S. Patent Nos. 5,350,674 and 5,585,362.

As discussed above, a "transfected", "or "transduced" host cell or cell line is one in which the genome has been altered or augmented by the presence of at least one heterologous or recombinant nucleic acid sequence. The host cells of the present invention are typically produced by transfection with a DNA sequence in a plasmid expression vector, a viral expression vector, or as an isolated linear DNA sequence. The transfected DNA can become a chromosomally integrated recombinant DNA sequence, which is composed of sequence encoding the siRNA.

To confirm the presence of the recombinant DNA sequence in the host cell, a variety of assays may be performed. Such assays include, for example, "molecular biological" assays well known to those of skill in the art, such as Southern and Northern blotting, RT-PCR and PCR; "biochemical" assays, such as detecting the presence or absence of a particular peptide, *e.g.*, by immunological means (ELISAs and Western blots) or by assays described herein to identify agents falling within the scope of the invention.

To detect and quantitate RNA produced from introduced recombinant DNA segments, RT-PCR may be employed. In this application of PCR, it is first necessary to reverse transcribe RNA into DNA, using enzymes such as

reverse transcriptase, and then through the use of conventional PCR techniques amplify the DNA. In most instances PCR techniques, while useful, will not demonstrate integrity of the RNA product. Further information about the nature of the RNA product may be obtained by Northern blotting. This technique demonstrates the presence of an RNA species and gives information about the integrity of that RNA. The presence or absence of an RNA species can also be determined using dot or slot blot Northern hybridizations. These techniques are modifications of Northern blotting and only demonstrate the presence or absence of an RNA species.

While Southern blotting and PCR may be used to detect the recombinant DNA segment in question, they do not provide information as to whether the preselected DNA segment is being expressed. Expression may be evaluated by specifically identifying the peptide products of the introduced recombinant DNA sequences or evaluating the phenotypic changes brought about by the expression of the introduced recombinant DNA segment in the host cell.

The instant invention provides a cell expression system for expressing exogenous nucleic acid material in a mammalian recipient. The expression system, also referred to as a "genetically modified cell", comprises a cell and an expression vector for expressing the exogenous nucleic acid material. The genetically modified cells are suitable for administration to a mammalian recipient, where they replace the endogenous cells of the recipient. Thus, the genetically modified cells are non-immortalized and are non-tumorigenic.

According to one embodiment, the cells are transfected or otherwise genetically modified *ex vivo*. The cells are isolated from a mammal (such as a human), nucleic acid introduced (*i.e.*, transduced or transfected *in vitro*) with a vector for expressing a heterologous (*e.g.*, recombinant) gene encoding the therapeutic agent, and then administered to a mammalian recipient for delivery of the therapeutic agent *in situ*. The mammalian recipient may be a human and the cells to be modified are autologous cells, *i.e.*, the cells are isolated from the mammalian recipient.

According to another embodiment, the cells are transfected or transduced or otherwise genetically modified *in vivo*. The cells from the mammalian recipient are transduced or transfected *in vivo* with a vector containing exogenous nucleic acid material for expressing a heterologous (*e.g.*, recombinant) gene encoding a therapeutic agent and the therapeutic agent is delivered *in situ*. As used herein, "exogenous nucleic acid material" refers to a nucleic acid or an oligonucleotide, either natural or synthetic, which is not naturally found in the cells; or if it is naturally found in the cells, is modified from its original or native form. Thus, "exogenous nucleic acid material" includes, for example, a non-naturally occurring nucleic acid that can be transcribed into an anti-sense RNA, a siRNA, as well as a "heterologous gene" (*i.e.*, a gene encoding a protein that is not expressed or is expressed at biologically insignificant levels in a naturally-occurring cell of the same type). To illustrate, a synthetic or natural gene encoding human erythropoietin (EPO) would be considered "exogenous nucleic acid material" with respect to human peritoneal mesothelial cells since the latter cells do not naturally express EPO. Still another example of "exogenous nucleic acid material" is the introduction of only part of a gene to create a recombinant gene, such as combining an regulatable promoter with an endogenous coding sequence via homologous recombination.

IV. Promoters of the Invention

As described herein, an expression cassette of the invention contains, *inter alia*, a promoter. Such promoters include the CMV promoter, as well as the RSV promoter, SV40 late promoter and retroviral LTRs (long terminal repeat elements), or brain cell specific promoters, although many other promoter elements well known to the art, such as tissue specific promoters or regulatable promoters may be employed in the practice of the invention.

In one embodiment of the present invention, an expression cassette may contain a pol II promoter that is operably linked to a nucleic acid sequence encoding a siRNA. Thus, the pol II promoter, *i.e.*, a RNA polymerase II dependent promoter, initiates the transcription of the siRNA. In another embodiment, the pol II promoter is regulatable.

Three RNA polymerases transcribe nuclear genes in eukaryotes. RNA polymerase II (pol II) synthesizes mRNA, *i.e.*, pol II transcribes the genes that encode proteins. In contrast, RNA polymerase I (pol I) and RNA polymerase III (pol III) transcribe only a limited set of transcripts, synthesizing RNAs that have structural or catalytic roles. RNA polymerase I makes the large ribosomal RNAs (rRNA), which are under the control of pol I promoters. RNA polymerase III makes a variety of small, stable RNAs, including the small 5S rRNA and transfer RNAs (tRNA), the transcription of which is under the control of pol III promoters.

As described herein, the inventors unexpectedly discovered that pol II promoters are useful to direct transcription of the siRNA. This was surprising because, as discussed above, pol II promoters are thought to be responsible for transcription of messenger RNA, *i.e.*, relatively long RNAs as compared to RNAs of 30 bases or less.

A pol II promoter may be used in its entirety, or a portion or fragment of the promoter sequence may be used in which the portion maintains the promoter activity. As discussed herein, pol II promoters are known to a skilled person in the art and include the promoter of any protein-encoding gene, *e.g.*, an endogenously regulated gene or a constitutively expressed gene. For example, the promoters of genes regulated by cellular physiological events, *e.g.*, heat shock, oxygen levels and/or carbon monoxide levels, *e.g.*, in hypoxia, may be used in the expression cassettes of the invention. In addition, the promoter of any gene regulated by the presence of a pharmacological agent, *e.g.*, tetracycline and derivatives thereof, as well as heavy metal ions and hormones may be employed in the expression cassettes of the invention. In an embodiment of the

invention, the pol II promoter can be the CMV promoter or the RSV promoter. In another embodiment, the pol II promoter is the CMV promoter.

As discussed above, a pol II promoter of the invention may be one naturally associated with an endogenously regulated gene or sequence, as may be obtained by isolating the 5′ non-coding sequences located upstream of the coding segment and/or exon. The pol II promoter of the expression cassette can be, for example, the same pol II promoter driving expression of the targeted gene of interest. Alternatively, the nucleic acid sequence encoding the siRNA may be placed under the control of a recombinant or heterologous pol II promoter, which refers to a promoter that is not normally associated with the targeted gene's natural environment. Such promoters include promoters isolated from any eukaryotic cell, and promoters not "naturally occurring," *i.e.,* containing different elements of different transcriptional regulatory regions, and/or mutations that alter expression. In addition to producing nucleic acid sequences of promoters synthetically, sequences may be produced using recombinant cloning and/or nucleic acid amplification technology, including PCR™, in connection with the compositions disclosed herein (see U.S. Patent 4,683,202, U.S. Patent 5,928,906, each incorporated herein by reference).

In one embodiment, a pol II promoter that effectively directs the expression of the siRNA in the cell type, organelle, and organism chosen for expression will be employed. Those of ordinary skill in the art of molecular biology generally know the use of promoters for protein expression, for example, see Sambrook and Russell (2001), incorporated herein by reference. The promoters employed may be constitutive, tissue-specific, inducible, and/or useful under the appropriate conditions to direct high level expression of the introduced DNA segment, such as is advantageous in the large-scale production of recombinant proteins and/or peptides. The identity of tissue-specific promoters, as well as assays to characterize their activity, is well known to those of ordinary skill in the art.

In another aspect of the invention, RNA molecules of the present invention can be expressed from transcription units (see for example Couture *et al.,* 1996,

TIG., 12, 510) inserted into DNA or RNA vectors. The recombinant vectors can be DNA plasmids or viral vectors. siRNA expressing viral vectors can be constructed based on, but not limited to, adeno-associated virus, retrovirus, adenovirus, or alphavirus. In another embodiment, pol III based constructs are used to express nucleic acid molecules of the invention (see for example Thompson, U.S. Pats. Nos. 5,902,880 and 6,146,886). The recombinant vectors capable of expressing the siRNA molecules can be delivered as described above, and persist in target cells. Alternatively, viral vectors can be used that provide for transient expression of nucleic acid molecules. Such vectors can be repeatedly administered as necessary. Once expressed, the siRNA molecule interacts with the target mRNA and generates an RNAi response. Delivery of siRNA molecule expressing vectors can be systemic, such as by intravenous or intra-muscular administration, by administration to target cells ex-planted from a subject followed by reintroduction into the subject, or by any other means that would allow for introduction into the desired target cell (for a review see Couture *et al.*, 1996, *TIG.*, 12, 510). In one aspect the invention features an expression vector comprising a nucleic acid sequence encoding at least one siRNA molecule of the instant invention. The expression vector can encode one or both strands of a siRNA duplex, or a single self-complementary strand that self hybridizes into a siRNA duplex. The nucleic acid sequences encoding the siRNA molecules of the instant invention can be operably linked in a manner that allows expression of the siRNA molecule (see for example Paul *et al.*, 2002, *Nature Biotechnology*, 19, 505; Miyagishi and Taira, 2002, *Nature Biotechnology*, 19, 497; Lee *et al.*, 2002, *Nature Biotechnology*, 19, 500; and Novina *et al.*, 2002, *Nature Medicine*, advance online publication doi:10.1038/nm725). In another aspect, the invention features an expression vector comprising: a) a transcription initiation region (*e.g.*, eukaryotic pol I, II or III initiation region); b) a transcription termination region (*e.g.*, eukaryotic pol I, II or III termination region); and c) a nucleic acid sequence encoding at least one of the siRNA molecules of the instant invention, wherein said sequence is operably linked to said initiation region and said termination region in a manner

that allows expression and/or delivery of the siRNA molecule. The vector can optionally include an open reading frame (ORF) for a protein operably linked on the 5' side or the 3'-side of the sequence encoding the siRNA of the invention; and/or an intron (intervening sequences).

Transcription of the siRNA molecule sequences can be driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters are expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type depends on the nature of the gene regulatory sequences (enhancers, silencers, etc.) present nearby. Prokaryotic RNA polymerase promoters are also used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990, *Proc. Natl. Acad. Sci. USA*, 87, 6743-7; Gao and Huang 1993, *Nucleic Acids Res.*, 21, 2867-72; Lieber *et al.*, 1993, *Methods Enzymol.*, 217, 47-66; Zhou *et al.*, 1990, *Mol. Cell. Biol.*, 10, 4529-37). Several investigators have demonstrated that nucleic acid molecules expressed from such promoters can function in mammalian cells (*e.g.* Kashani-Sabet *et al.*, 1992, *Antisense Res. Dev.*, 2, 3-15; Ojwang *et al.*, 1992, *Proc. Natl. Acad. Sci. USA*, 89, 10802-6; Chen *et al.*, 1992, *Nucleic Acids Res.*, 20, 4581-9; Yu *et al.*, 1993, *Proc. Natl. Acad. Sci. USA*, 90, 6340-4; L'Huillier *et al.*, 1992, *EMBO J.*, 11, 4411-8; Lisziewicz *et al.*, 1993, *Proc. Natl. Acad. Sci. U. S. A*, 90, 8000-4; Thompson *et al.*, 1995, *Nucleic Acids Res.*, 23, 2259; Sullenger & Cech, 1993, *Science*, 262, 1566). More specifically, transcription units such as the ones derived from genes encoding U6 small nuclear (snRNA), transfer RNA (tRNA) and adenovirus VA RNA are useful in generating high concentrations of desired RNA molecules such as siRNA in cells (Thompson *et al.*, *supra*; Couture and Stinchcomb, *1996, supra*; Noonberg *et al.*, 1994, *Nucleic Acid Res.*, 22, 2830; Noonberg *et al.*, U.S. Pat. No. 5,624,803; Good *et al.*, 1997, *Gene Ther.*, 4, 45; Beigelman *et al.*, International PCT Publication No. WO 96/18736. The above siRNA transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral

DNA vectors (such as adenovirus or adeno-associated virus vectors), or viral RNA vectors (such as retroviral or alphavirus vectors) (for a review see Couture and Stinchcomb, *1996, supra)*.

In another aspect the invention features an expression vector comprising a nucleic acid sequence encoding at least one of the siRNA molecules of the invention in a manner that allows expression of that siRNA molecule. The expression vector comprises in one embodiment; a) a transcription initiation region; b) a transcription termination region; and c) a nucleic acid sequence encoding at least one strand of the siRNA molecule, wherein the sequence is operably linked to the initiation region and the termination region in a manner that allows expression and/or delivery of the siRNA molecule.

In another embodiment the expression vector comprises: a) a transcription initiation region; b) a transcription termination region; c) an open reading frame; and d) a nucleic acid sequence encoding at least one strand of a siRNA molecule, wherein the sequence is operably linked to the 3'-end of the open reading frame and wherein the sequence is operably linked to the initiation region, the open reading frame and the termination region in a manner that allows expression and/or delivery of the siRNA molecule. In yet another embodiment, the expression vector comprises: a) a transcription initiation region; b) a transcription termination region; c) an intron; and d) a nucleic acid sequence encoding at least one siRNA molecule, wherein the sequence is operably linked to the initiation region, the intron and the termination region in a manner which allows expression and/or delivery of the nucleic acid molecule.

In another embodiment, the expression vector comprises: a) a transcription initiation region; b) a transcription termination region; c) an intron; d) an open reading frame; and e) a nucleic acid sequence encoding at least one strand of a siRNA molecule, wherein the sequence is operably linked to the 3'-end of the open reading frame and wherein the sequence is operably linked to the initiation region, the intron, the open reading frame and the termination region in a manner which allows expression and/or delivery of the siRNA molecule.

V. Methods for Introducing the Expression Cassettes of the Invention into Cells

The condition amenable to gene inhibition therapy may be a prophylactic process, *i.e.*, a process for preventing disease or an undesired medical condition. Thus, the instant invention embraces a system for delivering siRNA that has a prophylactic function (*i.e.*, a prophylactic agent) to the mammalian recipient.

The inhibitory nucleic acid material (*e.g.*, an expression cassette encoding siRNA directed to a gene of interest) can be introduced into the cell *ex vivo* or *in vivo* by genetic transfer methods, such as transfection or transduction, to provide a genetically modified cell. Various expression vectors (*i.e.*, vehicles for facilitating delivery of exogenous nucleic acid into a target cell) are known to one of ordinary skill in the art.

As used herein, "transfection of cells" refers to the acquisition by a cell of new nucleic acid material by incorporation of added DNA. Thus, transfection refers to the insertion of nucleic acid into a cell using physical or chemical methods. Several transfection techniques are known to those of ordinary skill in the art including: calcium phosphate DNA co-precipitation (Methods in Molecular Biology (1991)); DEAE-dextran (supra); electroporation (supra); cationic liposome-mediated transfection (supra); and tungsten particle-facilitated microparticle bombardment (Johnston (1990)). Strontium phosphate DNA co-precipitation (Brash *et al.* (1987)) is also a transfection method.

In contrast, "transduction of cells" refers to the process of transferring nucleic acid into a cell using a DNA or RNA virus. A RNA virus (*i.e.*, a retrovirus) for transferring a nucleic acid into a cell is referred to herein as a transducing chimeric retrovirus. Exogenous nucleic acid material contained within the retrovirus is incorporated into the genome of the transduced cell. A cell that has been transduced with a chimeric DNA virus (*e.g.*, an adenovirus carrying a cDNA encoding a therapeutic agent), will not have the exogenous nucleic acid material incorporated into its genome but will be capable of

expressing the exogenous nucleic acid material that is retained extrachromosomally within the cell.

The exogenous nucleic acid material can include the nucleic acid encoding the siRNA together with a promoter to control transcription. The promoter characteristically has a specific nucleotide sequence necessary to initiate transcription. The exogenous nucleic acid material may further include additional sequences (*i.e.*, enhancers) required to obtain the desired gene transcription activity. For the purpose of this discussion an "enhancer" is simply any non-translated DNA sequence that works with the coding sequence (in *cis*) to change the basal transcription level dictated by the promoter. The exogenous nucleic acid material may be introduced into the cell genome immediately downstream from the promoter so that the promoter and coding sequence are operatively linked so as to permit transcription of the coding sequence. An expression vector can include an exogenous promoter element to control transcription of the inserted exogenous gene. Such exogenous promoters include both constitutive and regulatable promoters.

Naturally-occurring constitutive promoters control the expression of essential cell functions. As a result, a nucleic acid sequence under the control of a constitutive promoter is expressed under all conditions of cell growth. Constitutive promoters include the promoters for the following genes which encode certain constitutive or "housekeeping" functions: hypoxanthine phosphoribosyl transferase (HPRT), dihydrofolate reductase (DHFR) (Scharfmann *et al.* (1991)), adenosine deaminase, phosphoglycerol kinase (PGK), pyruvate kinase, phosphoglycerol mutase, the beta-actin promoter (Lai *et al.* (1989)), and other constitutive promoters known to those of skill in the art. In addition, many viral promoters function constitutively in eukaryotic cells. These include: the early and late promoters of SV40; the long terminal repeats (LTRs) of Moloney Leukemia Virus and other retroviruses; and the thymidine kinase promoter of Herpes Simplex Virus, among many others.

Nucleic acid sequences that are under the control of regulatable promoters are expressed only or to a greater or lesser degree in the presence of

an inducing or repressing agent, (*e.g.*, transcription under control of the metallothionein promoter is greatly increased in presence of certain metal ions). Regulatable promoters include responsive elements (REs) that stimulate transcription when their inducing factors are bound. For example, there are REs for serum factors, steroid hormones, retinoic acid, cyclic AMP, and tetracycline and doxycycline. Promoters containing a particular RE can be chosen in order to obtain an regulatable response and in some cases, the RE itself may be attached to a different promoter, thereby conferring regulatability to the encoded nucleic acid sequence. Thus, by selecting the appropriate promoter (constitutive versus regulatable; strong versus weak), it is possible to control both the existence and level of expression of a nucleic acid sequence in the genetically modified cell. If the nucleic acid sequence is under the control of an regulatable promoter, delivery of the therapeutic agent *in situ* is triggered by exposing the genetically modified cell *in situ* to conditions for permitting transcription of the nucleic acid sequence, *e.g.*, by intraperitoneal injection of specific inducers of the regulatable promoters which control transcription of the agent. For example, *in situ* expression of a nucleic acid sequence under the control of the metallothionein promoter in genetically modified cells is enhanced by contacting the genetically modified cells with a solution containing the appropriate (*i.e.*, inducing) metal ions *in situ*.

Accordingly, the amount of siRNA generated *in situ* is regulated by controlling such factors as the nature of the promoter used to direct transcription of the nucleic acid sequence, (*i.e.*, whether the promoter is constitutive or regulatable, strong or weak) and the number of copies of the exogenous nucleic acid sequence encoding a siRNA sequence that are in the cell.

In addition to at least one promoter and at least one heterologous nucleic acid sequence encoding the siRNA, the expression vector may include a selection gene, for example, a neomycin resistance gene, for facilitating selection of cells that have been transfected or transduced with the expression vector.

Cells can also be transfected with two or more expression vectors, at least one vector containing the nucleic acid sequence(s) encoding the siRNA(s), the

other vector containing a selection gene. The selection of a suitable promoter, enhancer, selection gene and/or signal sequence is deemed to be within the scope of one of ordinary skill in the art without undue experimentation.

The following discussion is directed to various utilities of the instant invention. For example, the instant invention has utility as an expression system suitable for silencing the expression of gene(s) of interest.

The instant invention also provides various methods for making and using the above-described genetically-modified cells.

The instant invention also provides methods for genetically modifying cells of a mammalian recipient *in vivo*. According to one embodiment, the method comprises introducing an expression vector for expressing a siRNA sequence in cells of the mammalian recipient *in situ* by, for example, injecting the vector into the recipient.

VI. Delivery Vehicles for the Expression Cassettes of the Invention

Delivery of compounds into tissues and across the blood-brain barrier can be limited by the size and biochemical properties of the compounds. Currently, efficient delivery of compounds into cells *in vivo* can be achieved only when the molecules are small (usually less than 600 Daltons). Gene transfer for the correction of inborn errors of metabolism and neurodegenerative diseases of the central nervous system (CNS), and for the treatment of cancer has been accomplished with recombinant adenoviral vectors.

The selection and optimization of a particular expression vector for expressing a specific siRNA in a cell can be accomplished by obtaining the nucleic acid sequence of the siRNA, possibly with one or more appropriate control regions (*e.g.*, promoter, insertion sequence); preparing a vector construct comprising the vector into which is inserted the nucleic acid sequence encoding the siRNA; transfecting or transducing cultured cells *in vitro* with the vector construct; and determining whether the siRNA is present in the cultured cells.

Vectors for cell gene therapy include viruses, such as replication-deficient viruses (described in detail below). Exemplary viral vectors are derived from Harvey Sarcoma virus, ROUS Sarcoma virus, (MPSV), Moloney murine leukemia virus and DNA viruses (*e.g.*, adenovirus) (Ternin (1986)).

Replication-deficient retroviruses are capable of directing synthesis of all virion proteins, but are incapable of making infectious particles. Accordingly, these genetically altered retroviral expression vectors have general utility for high-efficiency transduction of nucleic acid sequences in cultured cells, and specific utility for use in the method of the present invention. Such retroviruses further have utility for the efficient transduction of nucleic acid sequences into cells *in vivo*. Retroviruses have been used extensively for transferring nucleic acid material into cells. Standard protocols for producing replication-deficient retroviruses (including the steps of incorporation of exogenous nucleic acid material into a plasmid, transfection of a packaging cell line with plasmid, production of recombinant retroviruses by the packaging cell line, collection of viral particles from tissue culture media, and infection of the target cells with the viral particles) are provided in Kriegler (1990) and Murray (1991).

An advantage of using retroviruses for gene therapy is that the viruses insert the nucleic acid sequence encoding the siRNA into the host cell genome, thereby permitting the nucleic acid sequence encoding the siRNA to be passed on to the progeny of the cell when it divides. Promoter sequences in the LTR region have been reported to enhance expression of an inserted coding sequence in a variety of cell types (see *e.g.*, Hilberg *et al.* (1987); Holland *et al.* (1987); Valerio *et al.* (1989). Some disadvantages of using a retrovirus expression vector are (1) insertional mutagenesis, *i.e.*, the insertion of the nucleic acid sequence encoding the siRNA into an undesirable position in the target cell genome which, for example, leads to unregulated cell growth and (2) the need for target cell proliferation in order for the nucleic acid sequence encoding the siRNA carried by the vector to be integrated into the target genome (Miller *et al.* (1990)).

Another viral candidate useful as an expression vector for transformation of cells is the adenovirus, a double-stranded DNA virus. The adenovirus is infective in a wide range of cell types, including, for example, muscle and endothelial cells (Larrick and Burck (1991)). The adenovirus also has been used as an expression vector in muscle cells *in vivo* (Quantin *et al.* (1992)).

Adenoviruses (Ad) are double-stranded linear DNA viruses with a 36 kb genome. Several features of adenovirus have made them useful as transgene delivery vehicles for therapeutic applications, such as facilitating *in vivo* gene delivery. Recombinant adenovirus vectors have been shown to be capable of efficient *in situ* gene transfer to parenchymal cells of various organs, including the lung, brain, pancreas, gallbladder, and liver. This has allowed the use of these vectors in methods for treating inherited genetic diseases, such as cystic fibrosis, where vectors may be delivered to a target organ. In addition, the ability of the adenovirus vector to accomplish *in situ* tumor transduction has allowed the development of a variety of anticancer gene therapy methods for non-disseminated disease. In these methods, vector containment favors tumor cell-specific transduction.

Like the retrovirus, the adenovirus genome is adaptable for use as an expression vector for gene therapy, *i.e.*, by removing the genetic information that controls production of the virus itself (Rosenfeld *et al.* (1991)). Because the adenovirus functions in an extrachromosomal fashion, the recombinant adenovirus does not have the theoretical problem of insertional mutagenesis.

Several approaches traditionally have been used to generate the recombinant adenoviruses. One approach involves direct ligation of restriction endonuclease fragments containing a nucleic acid sequence of interest to portions of the adenoviral genome. Alternatively, the nucleic acid sequence of interest may be inserted into a defective adenovirus by homologous recombination results. The desired recombinants are identified by screening individual plaques generated in a lawn of complementation cells.

Most adenovirus vectors are based on the adenovirus type 5 (Ad5) backbone in which an expression cassette containing the nucleic acid sequence

of interest has been introduced in place of the early region 1 (E1) or early region 3 (E3). Viruses in which E1 has been deleted are defective for replication and are propagated in human complementation cells (*e.g.*, 293 or 911 cells), which supply the missing gene E1 and pIX in *trans*.

In one embodiment of the present invention, one will desire to generate siRNA in a brain cell or brain tissue. A suitable vector for this application is an FIV vector (Brooks *et al.* (2002); Alisky *et al.* (2000a)) or an AAV vector. For example, one may use AAV5 (Davidson *et al.* (2000); Alisky *et al.* (2000a)). Also, one may apply poliovirus (Bledsoe *et al.* (2000)) or HSV vectors (Alisky *et al.* (2000b)).

Adeno associated virus (AAV) is a small nonpathogenic virus of the parvoviridae family (for review see Muzyczka, N. 1992. Curr Top Microbiol Immunol 158: 97-129; see also U.S. Patent No. 6,468,524). AAV is distinct from the other members of this family by its dependence upon a helper virus for replication. In the absence of a helper virus, AAV may integrate in a locus specific manner into the q arm of chromosome 19 (Kotin *et al.*, (1990) Proc. Natl. Acad. Sci. (USA) 87: 2211-2215). The approximately 5 kb genome of AAV consists of one segment of single stranded DNA of either plus or minus polarity. The ends of the genome are short inverted terminal repeats which can fold into hairpin structures and serve as the origin of viral DNA replication. Physically, the parvovirus virion is non-enveloped and its icosohedral capsid is approximately 20 nm in diameter.

To-date seven serologically distinct AAVs have been identified and five have been isolated from humans or primates and are referred to as AAV types 1-5 (Arella et al Handbook of Parvoviruses. Vol. 1. ed. P. Tijssen. Boca Raton, Fla., CRC Press, 1990). The most extensively studied of these isolates is AAV type 2 (AAV2). The genome of AAV2 is 4680 nucleotides in length and contains two open reading frames (ORFs). The left ORF encodes the non-structural Rep proteins, Rep40, Rep 52, Rep68 and Rep 78, which are involved in regulation of replication and transcription in addition to the production of single-stranded progeny genomes. Furthermore, two of the Rep proteins have

been associated with the possible integration of AAV genomes into a region of the q arm of human chromosome 19. Rep68/78 have also been shown to possess NTP binding activity as well as DNA and RNA helicase activities. The Rep proteins possess a nuclear localization signal as well as several potential phosphorylation sites. Mutation of one of these kinase sites resulted in a loss of replication activity.

The ends of the genome are short inverted terminal repeats which have the potential to fold into T-shaped hairpin structures that serve as the origin of viral DNA replication. Within the ITR region two elements have been described which are central to the function of the ITR, a GAGC repeat motif and the terminal resolution site (trs). The repeat motif has been shown to bind Rep when the ITR is in either a linear or hairpin conformation. This binding serves to position Rep68/78 for cleavage at the trs which occurs in a site- and strand-specific manner. In addition to their role in replication, these two elements appear to be central to viral integration. Contained within the chromosome 19 integration locus is a Rep binding site with an adjacent trs. These elements have been shown to be functional and necessary for locus specific integration.

The AAV2 virion is a non-enveloped, icosohedral particle approximately 25 nm in diameter, consisting of three related proteins referred to as VPI,2 and 3. The right ORF encodes the capsid proteins, VP1, VP2, and VP3. These proteins are found in a ratio of 1:1:10 respectively and are all derived from the right-hand ORF. The capsid proteins differ from each other by the use of alternative splicing and an unusual start codon. Deletion analysis has shown that removal or alteration of VP1 which is translated from an alternatively spliced message results in a reduced yield of infections particles. Mutations within the VP3 coding region result in the failure to produce any single-stranded progeny DNA or infectious particles.

The following features of AAV have made it an attractive vector for gene transfer. AAV vectors have been shown *in vitro* to stably integrate into the cellular genome; possess a broad host range; transduce both dividing and non dividing cells *in vitro* and *in vivo* and maintain high levels of expression of the

transduced genes. Viral particles are heat stable, resistant to solvents, detergents, changes in pH, temperature, and can be concentrated on CsCl gradients. Integration of AAV provirus is not associated with any long term negative effects on cell growth or differentiation. The ITRs have been shown to be the only cis elements required for replication, packaging and integration and may contain some promoter activities.

Further provided by this invention are chimeric viruses where AAV can be combined with herpes virus, herpes virus amplicons, baculovirus or other viruses to achieve a desired tropism associated with another virus. For example, the AAV4 ITRs could be inserted in the herpes virus and cells could be infected. Post-infection, the ITRs of AAV4 could be acted on by AAV4 rep provided in the system or in a separate vehicle to rescue AAV4 from the genome. Therefore, the cellular tropism of the herpes simplex virus can be combined with AAV4 rep mediated targeted integration. Other viruses that could be utilized to construct chimeric viruses include lentivirus, retrovirus, pseudotyped retroviral vectors, and adenoviral vectors.

Also provided by this invention are variant AAV vectors. For example, the sequence of a native AAV, such as AAV5, can be modified at individual nucleotides. The present invention includes native and mutant AAV vectors. The present invention further includes all AAV serotypes.

Thus, as will be apparent to one of ordinary skill in the art, a variety of suitable viral expression vectors are available for transferring exogenous nucleic acid material into cells. The selection of an appropriate expression vector to express a therapeutic agent for a particular condition amenable to gene silencing therapy and the optimization of the conditions for insertion of the selected expression vector into the cell, are within the scope of one of ordinary skill in the art without the need for undue experimentation.

In another embodiment, the expression vector is in the form of a plasmid, which is transferred into the target cells by one of a variety of methods: physical (*e.g.*, microinjection (Capecchi (1980)), electroporation (Andreason and Evans (1988), scrape loading, microparticle bombardment (Johnston (1990)) or by

cellular uptake as a chemical complex (*e.g.*, calcium or strontium co-precipitation, complexation with lipid, complexation with ligand) (Methods in Molecular Biology (1991)). Several commercial products are available for cationic liposome complexation including Lipofectin™ (Gibco-BRL, Gaithersburg, Md.) (Felgner *et al.* (1987)) and Transfectam™ (Promega, Madison, Wis.) (Behr *et al.* (1989); Loeffler *et al.* (1990)). However, the efficiency of transfection by these methods is highly dependent on the nature of the target cell and accordingly, the conditions for optimal transfection of nucleic acids into cells using the above-mentioned procedures must be optimized. Such optimization is within the scope of one of ordinary skill in the art without the need for undue experimentation.

VII. Diseases and Conditions Amendable to the Methods of the Invention

In the certain embodiments of the present invention, a mammalian recipient to an expression cassette of the invention has a condition that is amenable to gene silencing therapy. As used herein, "gene silencing therapy" refers to administration to the recipient exogenous nucleic acid material encoding a therapeutic siRNA and subsequent expression of the administered nucleic acid material *in situ*. Thus, the phrase "condition amenable to siRNA therapy" embraces conditions such as genetic diseases (*i.e.*, a disease condition that is attributable to one or more gene defects), acquired pathologies (*i.e.*, a pathological condition that is not attributable to an inborn defect), cancers, neurodegenerative diseases, *e.g.*, trinucleotide repeat disorders, and prophylactic processes (*i.e.*, prevention of a disease or of an undesired medical condition). A gene "associated with a condition" is a gene that is either the cause, or is part of the cause, of the condition to be treated. Examples of such genes include genes associated with a neurodegenerative disease (*e.g.*, a trinucleotide-repeat disease such as a disease associated with polyglutamine repeats, Huntington's disease, and several spinocerebellar ataxias), and genes encoding ligands for chemokines involved in the migration of a cancer cells, or chemokine receptor. Also siRNA

expressed from viral vectors may be used for *in vivo* antiviral therapy using the vector systems described.

Accordingly, as used herein, the term "therapeutic siRNA" refers to any siRNA that has a beneficial effect on the recipient. Thus, "therapeutic siRNA" embraces both therapeutic and prophylactic siRNA.

Differences between alleles that are amenable to targeting by siRNA include disease-causing mutations as well as polymorphisms that are not themselves mutations, but may be linked to a mutation or associated with a predisposition to a disease state. Examples of targetable disease mutations include tau mutations that cause frontotemporal dementia and the GAG deletion in the TOR1A gene that causes DYT1 dystonia. An example of a targetable polymorphism that is not itself a mutation is the C/G single nucleotide polymorphism (G987C) in the MJD1 gene immediately downstream of the mutation that causes spinocerebellar ataxia type 3 and the polymorphism in exon 58 associated with Huntington's disease.

Single nucleotide polymorphisms comprise most of the genetic diversity between humans, and that many disease genes, including the HD gene in Huntington's disease, contain numerous single nucleotide or multiple nucleotide polymorphisms that could be separately targeted in one allele vs. the other, as shown in Figure 15. The major risk factor for developing Alzheimer's disease is the presence of a particular polymorphism in the apolipoprotein E gene.

A. Gene defects

A number of diseases caused by gene defects have been identified. For example, this strategy can be applied to a major class of disabling neurological disorders. For example this strategy can be applied to the polyglutamine diseases, as is demonstrated by the reduction of polyglutamine aggregation in cells following application of the strategy. The neurodegenerative disease may be a trinucleotide-repeat disease, such as a disease associated with polyglutamine repeats, including Huntington's disease, and several spinocerebellar ataxias.

Additionally, this strategy can be applied to a non-degenerative neurological disorder, such as DYT1 dystonia.

B. Acquired pathologies

As used herein, "acquired pathology" refers to a disease or syndrome manifested by an abnormal physiological, biochemical, cellular, structural, or molecular biological state. For example, the disease could be a viral disease, such as hepatitis or AIDS.

C. Cancers

The condition amenable to gene silencing therapy alternatively can be a genetic disorder or an acquired pathology that is manifested by abnormal cell proliferation, *e.g.*, cancer. According to this embodiment, the instant invention is useful for silencing a gene involved in neoplastic activity. The present invention can also be used to inhibit overexpression of one or several genes. The present invention can be used to treat neuroblastoma, medulloblastoma, or glioblastoma.

VIII. Dosages, Formulations and Routes of Administration of the Agents of the Invention

The agents of the invention are administered so as to result in a reduction in at least one symptom associated with a disease. The amount administered will vary depending on various factors including, but not limited to, the composition chosen, the particular disease, the weight, the physical condition, and the age of the mammal, and whether prevention or treatment is to be achieved. Such factors can be readily determined by the clinician employing animal models or other test systems which are well known to the art.

Administration of siRNA may be accomplished through the administration of the nucleic acid molecule encoding the siRNA (see, for example, Felgner *et al.*, U.S. Patent No. 5,580,859, Pardoll *et al.* 1995; Stevenson *et al.* 1995; Molling 1997; Donnelly *et al.* 1995; Yang *et al.* II; Abdallah *et al.* 1995). Pharmaceutical formulations, dosages and routes of

administration for nucleic acids are generally disclosed, for example, in Felgner *et al., supra.*

The present invention envisions treating a disease, for example, a neurodegenerative disease, in a mammal by the administration of an agent, *e.g.*, a nucleic acid composition, an expression vector, or a viral particle of the invention. Administration of the therapeutic agents in accordance with the present invention may be continuous or intermittent, depending, for example, upon the recipient's physiological condition, whether the purpose of the administration is therapeutic or prophylactic, and other factors known to skilled practitioners. The administration of the agents of the invention may be essentially continuous over a preselected period of time or may be in a series of spaced doses. Both local and systemic administration is contemplated.

One or more suitable unit dosage forms having the therapeutic agent(s) of the invention, which, as discussed below, may optionally be formulated for sustained release (for example using microencapsulation, see WO 94/07529, and U.S. Patent No. 4,962,091 the disclosures of which are incorporated by reference herein), can be administered by a variety of routes including parenteral, including by intravenous and intramuscular routes, as well as by direct injection into the diseased tissue. For example, the therapeutic agent may be directly injected into the brain. Alternatively the therapeutic agent may be introduced intrathecally for brain and spinal cord conditions. In another example, the therapeutic agent may be introduced intramuscularly for viruses that traffic back to affected neurons from muscle, such as AAV, lentivirus and adenovirus. The formulations may, where appropriate, be conveniently presented in discrete unit dosage forms and may be prepared by any of the methods well known to pharmacy. Such methods may include the step of bringing into association the therapeutic agent with liquid carriers, solid matrices, semi-solid carriers, finely divided solid carriers or combinations thereof, and then, if necessary, introducing or shaping the product into the desired delivery system.

When the therapeutic agents of the invention are prepared for administration, they may be combined with a pharmaceutically acceptable

carrier, diluent or excipient to form a pharmaceutical formulation, or unit dosage form. The total active ingredients in such formulations include from 0.1 to 99.9% by weight of the formulation. A "pharmaceutically acceptable" is a carrier, diluent, excipient, and/or salt that is compatible with the other ingredients of the formulation, and not deleterious to the recipient thereof. The active ingredient for administration may be present as a powder or as granules; as a solution, a suspension or an emulsion.

Pharmaceutical formulations containing the therapeutic agents of the invention can be prepared by procedures known in the art using well known and readily available ingredients. The therapeutic agents of the invention can also be formulated as solutions appropriate for parenteral administration, for instance by intramuscular, subcutaneous or intravenous routes.

The pharmaceutical formulations of the therapeutic agents of the invention can also take the form of an aqueous or anhydrous solution or dispersion, or alternatively the form of an emulsion or suspension.

Thus, the therapeutic agent may be formulated for parenteral administration (*e.g.*, by injection, for example, bolus injection or continuous infusion) and may be presented in unit dose form in ampules, pre-filled syringes, small volume infusion containers or in multi-dose containers with an added preservative. The active ingredients may take such forms as suspensions, solutions, or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredients may be in powder form, obtained by aseptic isolation of sterile solid or by lyophilization from solution, for constitution with a suitable vehicle, *e.g.*, sterile, pyrogen-free water, before use.

It will be appreciated that the unit content of active ingredient or ingredients contained in an individual aerosol dose of each dosage form need not in itself constitute an effective amount for treating the particular indication or disease since the necessary effective amount can be reached by administration of a plurality of dosage units. Moreover, the effective amount may be achieved

using less than the dose in the dosage form, either individually, or in a series of administrations.

The pharmaceutical formulations of the present invention may include, as optional ingredients, pharmaceutically acceptable carriers, diluents, solubilizing or emulsifying agents, and salts of the type that are well-known in the art. Specific non-limiting examples of the carriers and/or diluents that are useful in the pharmaceutical formulations of the present invention include water and physiologically acceptable buffered saline solutions such as phosphate buffered saline solutions pH 7.0-8.0. saline solutions and water.

The invention will now be illustrated by the following non-limiting Example.

Example 1

siRNA-Mediated Silencing of Genes Using Viral Vectors

In this Example, it is shown that genes can be silenced in an allele-specific manner. It is also demonstrated that viral-mediated delivery of siRNA can specifically reduce expression of targeted genes in various cell types, both *in vitro* and *in vivo*. This strategy was then applied to reduce expression of a neurotoxic polyglutamine disease protein. The ability of viral vectors to transduce cells efficiently *in vivo*, coupled with the efficacy of virally expressed siRNA shown here, extends the application of siRNA to viral-based therapies and *in vivo* targeting experiments that aim to define the function of specific genes.

Experimental Protocols

Generation of the expression cassettes and viral vectors. The modified CMV (mCMV) promoter was made by PCR amplification of CMV by primers

5'-AAGGTACCAGATCTTAGTTATTAATAGTAATCAATTACGG-3' (SEQ ID NO:1) and

5'-GAATCGATGCATGCCTCGAGACGGTTCACTAAACCAGCTCTGC-3' (SEQ ID NO:2) with peGFPN1 plasmid (purchased from Clontech, Inc) as template. The mCMV product was cloned into the Kpn*I* and Cla*I* sites of the adenoviral shuttle vector pAd5KnpA, and was named pmCMVknpA. To construct the minimal polyA cassette, the oligonucleotides, 5'-CTAGAACTAGTAATAAAGGATCCTTTATTTTCATTGGATCCGTGTGTT GGTTTTTTGTGTGCGGCCGCG-3' (SEQ ID NO:3) and 5'-TCGACGCGGCCGCACACAAAAAACCAACACACGGATCC AATGAAAATAAAGGATCCTTTATTACTAGTT-3' (SEQ ID NO:4), were used. The oligonucleotides contain Spe*I* and Sal*I* sites at the 5' and 3' ends, respectively. The synthesized polyA cassette was ligated into Spe*I*, Sal*I* digested pmCMVKnpA. The resultant shuttle plasmid, pmCMVmpA was used for construction of head-to-head 21bp hairpins of eGFP (bp 418 to 438), human β-glucuronidase (bp 649 to 669), mouse β-glucuronidase (bp 646 to 666) or *E. coli* β-galactosidase (bp 1152-1172). The eGFP hairpins were also cloned into the Ad shuttle plasmid containing the commercially available CMV promoter and polyA cassette from SV40 large T antigen (pCMVsiGFPx). Shuttle plasmids were co-transfected into HEK293 cells along with the adenovirus backbones for generation of full-length Ad genomes. Viruses were harvested 6-10 days after transfection and amplified and purified as described (Anderson, R.D., *et al.*, *Gene Ther.* 7:1034-1038 (2000)).

Northern blotting. Total RNA was isolated from HEK293 cells transfected by plasmids or infected by adenoviruses using TRIZOL®Reagent (Invitrogen™ Life Technologies, Carlsbad, CA) according to the manufacturer's instruction. RNAs (30μg) were separated by electrophoresis on 15% (wt/vol) polyacrylamide-urea gels to detect transcripts, or on 1% agarose-formaldehyde gel for target mRNAs analysis. RNAs were transferred by electroblotting onto hybond N+ membrane (Amersham Pharmacia Biotech). Blots were probed with ^{32}P-labeled sense (5'-CACAAGCTGGAGTACAACTAC-3' (SEQ ID NO:5)) or antisense (5'-GTACTTGTACTCCAGCTTTGTG-3' (SEQ ID NO:6)) oligonucleotides at 37°C for 3h for evaluation of siRNA transcripts, or probed

for target mRNAs at 42°C overnight. Blots were washed using standard methods and exposed to film overnight. *In vitro* studies were performed in triplicate with a minimum of two repeats.

***In vivo* studies and tissue analyses.** All animal procedures were approved by the University of Iowa Committee on the Care and Use of Animals. Mice were injected into the tail vein (n = 10 per group) or into the brain (n = 6 per group) as described previously (Stein, C.S., *et al.*, *J. Virol.* 73:3424-3429 (1999)) with the virus doses indicated. Animals were sacrificed at the noted times and tissues harvested and sections or tissue lysates evaluated for β-glucuronidase expression, eGFP fluorescence, or β-galactosidase activity using established methods (Xia, H. *et al.*, *Nat. Biotechnol.* 19:640-644 (2001)). Total RNA was harvested from transduced liver using the methods described above.

Cell Lines. PC12 tet off cell lines (Clontech Inc., Palo Alto, CA) were stably transfected with a tetracycline regulatable plasmid into which was cloned GFPQ19 or GFPQ80 (Chai, Y. *et al.*, *J. Neurosci.* 19:10338-10347 (1999)). For GFP-Q80, clones were selected and clone 29 chosen for regulatable properties and inclusion formation. For GFP-Q19 clone 15 was selected for uniformity of GFP expression following gene expression induction. In all studies 1.5 μg/ml dox was used to repress transcription. All experiments were done in triplicate and were repeated 4 times.

Results and Discussion

To accomplish intracellular expression of siRNA, a 21-bp hairpin representing sequences directed against eGFP was constructed, and its ability to reduce target gene expression in mammalian cells using two distinct constructs was tested. Initially, the siRNA hairpin targeted against eGFP was placed under the control of the CMV promoter and contained a full-length SV-40 polyadenylation (polyA) cassette (pCMVsiGFPx). In the second construct, the hairpin was juxtaposed almost immediate to the CMV transcription start site (within 6 bp) and was followed by a synthetic, minimal polyA cassette (Fig. 1A, pmCMVsiGFPmpA) (Experimental Protocols), because we reasoned that

functional siRNA would require minimal to no overhangs (Caplan, N.J., *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* 98:9742-9747 (2001); Nykänen, A., *et al.*, *Cell* 107:309-321 (2001)). Co-transfection of pmCMVsiGFPmpA with pEGFPN1 (Clontech Inc) into HEK293 cells markedly reduced eGFP fluorescence (Fig. 1C). pmCMVsiGFPmpA transfection led to the production of an approximately 63 bp RNA specific for eGFP (Fig. 1D), consistent with the predicted size of the siGFP hairpin-containing transcript. Reduction of target mRNA and eGFP protein expression was noted in pmCMVsiGFPmpA-transfected cells only (Fig. 1E, F). In contrast, eGFP RNA, protein and fluorescence levels remained unchanged in cells transfected with pEGFPN1 and pCMVsiGFPx (Fig. 1E, G), pEGFPN1 and pCMVsißglucmpA (Fig. 1E, F, H), or pEGFPN1 and pCMVsißgalmpA, the latter expressing siRNA against *E. coli* ß-galactosidase (Fig. 1E). These data demonstrate the specificity of the expressed siRNAs.

Constructs identical to pmCMVsiGFPmpA except that a spacer of 9, 12 and 21 nucleotides was present between the transcription start site and the 21 bp hairpin were also tested. In each case, there was no silencing of eGFP expression (data not shown). Together the results indicate that the spacing of the hairpin immediate to the promoter can be important for functional target reduction, a fact supported by recent studies in MCF-7 cells (Brummelkamp, T.R., *et al.*, *Science* 296:550-553 (2002)).

Recombinant adenoviruses were generated from the siGFP (pmCMVsiGFPmpA) and siβgluc (pmCMVsiβglucmpA) plasmids (Xia, H., *et al.*, *Nat. Biotechnol.* 19:640-644 (2001); Anderson, R.D., *et al.*, *Gene Ther.* 7:1034-1038 (2000)) to test the hypothesis that virally expressed siRNA allows for diminished gene expression of endogenous targets *in vitro* and *in vivo*. HeLa cells are of human origin and contain moderate levels of the soluble lysosomal enzyme β-glucuronidase. Infection of HeLa cells with viruses expressing siβgluc caused a specific reduction in human ß-glucuronidase mRNA (Fig. 1I) leading to a 60% decrease in β-glucuronidase activity relative to siGFP or control cells (Fig 1J). Optimization of siRNA sequences using methods to refine target mRNA accessible sequences (Lee, N.S., *et al.*, *Nat. Biotechnol.* 19:500-

505 (2002)) could improve further the diminution of ß-glucuronidase transcript and protein levels.

The results in Fig. 1 are consistent with earlier work demonstrating the ability of synthetic 21-bp double stranded RNAs to reduce expression of target genes in mammalian cells following transfection, with the important difference that in the present studies the siRNA was synthesized intracellularly from readily available promoter constructs. The data support the utility of regulatable, tissue or cell-specific promoters for expression of siRNA when suitably modified for close juxtaposition of the hairpin to the transcriptional start site and inclusion of the minimal polyA sequence containing cassette (*see*, Methods above).

To evaluate the ability of virally expressed siRNA to diminish target-gene expression in adult mouse tissues *in vivo*, transgenic mice expressing eGFP (Okabe, M. *et al.*, *FEBS Lett.* 407:313-319 (1997)) were injected into the striatal region of the brain with 1 x 10^7 infectious units of recombinant adenovirus vectors expressing siGFP or control siβgluc. Viruses also contained a dsRed expression cassette in a distant region of the virus for unequivocal localization of the injection site. Brain sections evaluated 5 days after injection by fluorescence (Fig. 2A) or western blot assay (Fig. 2B) demonstrated reduced eGFP expression. Decreased eGFP expression was confined to the injected hemisphere (Fig. 2B). The *in vivo* reduction is promising, particularly since transgenically expressed eGFP is a stable protein, making complete reduction in this short time frame unlikely. Moreover, evaluation of eGFP levels was done 5 days after injection, when inflammatory changes induced by the adenovirus vector likely enhance transgenic eGFP expression from the CMV enhancer (Ooboshi, H., *et al.*, *Arterioscler. Thromb. Vasc. Biol.* 17:1786-1792 (1997)).

It was next tested whether virus mediated siRNA could decrease expression from endogenous alleles *in vivo*. Its ability to decrease β-glucuronidase activity in the murine liver, where endogenous levels of this relatively stable protein are high, was evaluated. Mice were injected via the tail vein with a construct expressing murine-specific siβgluc (AdsiMuβgluc), or the control viruses Adsiβgluc (specific for human β-glucuronidase) or Adsiβgal.

Adenoviruses injected into the tail vein transduced hepatocytes as shown previously (Stein, C.S., *et al.*, *J. Virol.* 73:3424-3429 (1999)). Liver tissue harvested 3 days later showed specific reduction of target ß-glucuronidase RNA in AdsiMußgluc treated mice only (Fig. 2C). Fluorometric enzyme assay of liver lysates confirmed these results, with a 12% decrease in activity from liver harvested from AdsiMuβgluc injected mice relative to Adsiβgal and Adsiβgluc treated ones ($p<0.01$; $n=10$). Interestingly, sequence differences between the murine and human siRNA constructs are limited, with 14 of 21 bp being identical. These results confirm the specificity of virus mediated siRNA, and indicate that allele-specific applications are possible. Together, the data are the first to demonstrate the utility of siRNA to diminish target gene expression in brain and liver tissue *in vivo*, and establish that allele-specific silencing *in vivo* is possible with siRNA.

One powerful therapeutic application of siRNA is to reduce expression of toxic gene products in dominantly inherited diseases such as the polyglutamine (polyQ) neurodegenerative disorders (Margolis, R.L. & Ross, C.A. *Trends Mol. Med.* 7:479-482 (2001)). The molecular basis of polyQ diseases is a novel toxic property conferred upon the mutant protein by polyQ expansion. This toxic property is associated with disease protein aggregation. The ability of virally expressed siRNA to diminish expanded polyQ protein expression in neural PC-12 clonal cell lines was evaluated. Lines were developed that express tetracycline-repressible eGFP-polyglutamine fusion proteins with normal or expanded glutamine of 19 (eGFP-Q19) and 80 (eGFP-Q80) repeats, respectively. Differentiated, eGFP-Q19-expressing PC12 neural cells infected with recombinant adenovirus expressing siGFP demonstrated a specific and dose-dependent decrease in eGFP-Q19 fluorescence (Fig. 3A, C) and protein levels (Fig. 3B). Application of Adsiβgluc as a control had no effect (Fig. 3A-C). Quantitative image analysis of eGFP fluorescence demonstrated that siGFP reduced GFPQ19 expression by greater than 96% and 93% for 100 and 50 MOI respectively, relative to control siRNA (Fig. 3C). The multiplicity of infection (MOI) of 100 required to achieve maximal inhibition of eGFP-Q19 expression

results largely from the inability of PC12 cells to be infected by adenovirus-based vectors. This barrier can be overcome using AAV- or lentivirus-based expression systems (Davidson, B.L., *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* 97:3428-3432 (2000); Brooks, A.I., et al, *Proc. Natl. Acad. Sci. U. S. A.* 99:6216-6221 (2002)).

To test the impact of siRNA on the size and number of aggregates formed in eGFP-Q80 expressing cells, differentiated PC-12/eGFP-Q80 neural cells were infected with AdsiGFP or Adsiβgluc 3 days after doxycycline removal to induce GFP-Q80 expression. Cells were evaluated 3 days later. In mock-infected control cells (Fig. 4A), aggregates were very large 6 days after induction as reported by others (Chai, Y., *et al.*, *J. Neurosci.* 19:10338-10347 (1999; Moulder, K.L., *et al.*, *J. Neurosci.* 19:705-715 (1999)). Large aggregates were also seen in cells infected with Adsiβgluc (Fig. 4B), AdsiGFPx, (Fig. 4C, siRNA expressed from the normal CMV promoter and containing the SV40 large T antigen polyadenylation cassette), or Adsiβgal (Fig. 4D). In contrast, polyQ aggregate formation was significantly reduced in AdsiGFP infected cells (Fig. 4E), with fewer and smaller inclusions and more diffuse eGFP fluorescence. AdsiGFP-mediated reduction in aggregated and monomeric GFP-Q80 was verified by Western blot analysis (Fig. 4F), and quantitation of cellular fluorescence (Fig. 4G). AdsiGFP caused a dramatic and specific, dose-dependent reduction in eGFP-Q80 expression (Fig. 4F, G).

It was found that transcripts expressed from the modified CMV promoter and containing the minimal polyA cassette were capable of reducing gene expression in both plasmid and viral vector systems (Figs. 1-4). The placement of the hairpin immediate to the transcription start site and use of the minimal polyadenylation cassette was of critical importance. In plants and Drosophila, RNA interference is initiated by the ATP-dependent, processive cleavage of long dsRNA into 21-25 bp double-stranded siRNA, followed by incorporation of siRNA into a RNA-induced silencing complex that recognizes and cleaves the target (Nykänen, A., *et al.*, *Cell* 107:309-321 (2001); Zamore, PD., *et al.*, *Cell* 101:25-33 (2000); Bernstein, E., *et al.*, *Nature* 409:363-366 (2001); Hamilton,

A.J. & Baulcombe, D.C. *Science* 286:950-952 (1999); Hammond, S.M. *et al.*, *Nature* 404:293-296 (2000)). Viral vectors expressing siRNA are useful in determining if similar mechanisms are involved in target RNA cleavage in mammalian cells *in vivo*.

In summary, these data demonstrate that siRNA expressed from viral vectors *in vitro* and *in vivo* specifically reduce expression of stably expressed plasmids in cells, and endogenous transgenic targets in mice. Importantly, the application of virally expressed siRNA to various target alleles in different cells and tissues *in vitro* and *in vivo* was demonstrated. Finally, the results show that it is possible to reduce polyglutamine protein levels in neurons, which is the cause of at least nine inherited neurodegenerative diseases, with a corresponding decrease in disease protein aggregation. The ability of viral vectors based on adeno-associated virus (Davidson, B.L., *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* 97:3428-3432 (2000)) and lentiviruses (Brooks, A.I., *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* 99:6216-6221 (2002)) to efficiently transduce cells in the CNS, coupled with the effectiveness of virally-expressed siRNA demonstrated here, extends the application of siRNA to viral-based therapies and to basic research, including inhibiting novel ESTs to define gene function.

Example 2

siRNA Suppression of Genes Involved in MJD/SCA3 and FTDP-17

Modulation of gene expression by endogenous, noncoding RNAs is increasingly appreciated to play a role in eukaryotic development, maintenance of chromatin structure and genomic integrity. Recently, techniques have been developed to trigger RNA interference (RNAi) against specific targets in mammalian cells by introducing exogenously produced or intracellularly expressed siRNAs. These methods have proven to be quick, inexpensive and effective for knockdown experiments *in vitro* and *in vivo*. The ability to accomplish selective gene silencing has led to the hypothesis that siRNAs might be employed to suppress gene expression for therapeutic benefit.

Dominantly inherited diseases are ideal candidates for siRNA-based therapy. To explore the utility of siRNA in inherited human disorders, the inventors employed cellular models to test whether we could target mutant alleles causing two classes of dominantly inherited, untreatable neurodegenerative diseases: polyglutamine (polyQ) neurodegeneration in MJD/SCA3 and frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17). The polyQ neurodegenerative disorders consist of at least nine diseases caused by CAG repeat expansions that encode polyQ in the disease protein. PolyQ expansion confers a dominant toxic property on the mutant protein that is associated with aberrant accumulation of the disease protein in neurons. In FTDP-17, Tau mutations lead to the formation of neurofibrillary tangles accompanied by neuronal dysfunction and degeneration. The precise mechanisms by which these mutant proteins cause neuronal injury are unknown, but considerable evidence suggests that the abnormal proteins themselves initiate the pathogenic process. Accordingly, eliminating expression of the mutant protein by siRNA or other means should, in principle, slow or even prevent disease. However, because many dominant disease genes may also encode essential proteins, the inventors sought to develop siRNA-mediated approaches that selectively inactivate mutant alleles while allowing continued expression of the wild type protein.

Methods

siRNA Synthesis. *In vitro* siRNA synthesis was previously described (Donze 2000). Reactions were performed with desalted DNA oligonucleotides (IDT Coralville, IA) and the AmpliScribeT7 High Yield Transcription Kit (Epicentre Madison, WI). Yield was determined by absorbance at 260nm. Annealed siRNAs were assessed for double stranded character by agarose gel (1% w/v) electrophoresis and ethidium bromide staining. Note that for all siRNAs generated in this study the most 5' nucleotide in the targeted cDNA sequence is referred to as position 1 and each subsequent nucleotide is numbered in ascending order from 5' to 3'.

Plasmid Construction. The human ataxin-3 cDNA was expanded to 166 CAG's by PCR (Laccone 1999). PCR products were digested at BamHI and KpnI sites introduced during PCR and ligated into BglII and KpnI sites of pEGFP-N1 (Clontech) resulting in full-length expanded ataxin-3 fused to the N-terminus of EGFP. Untagged Ataxin-3-Q166 was constructed by ligating a PpuMI-NotI ataxin-3 fragment (3' of the CAG repeat) into Ataxin-3-Q166-GFP cut with PpuMI and NotI to remove EGFP and replace the normal ataxin-3 stop codon. Ataxin-3-Q28-GFP was generated as above from pcDNA3.1-ataxin-3-Q28. Constructs were sequence verified to ensure that no PCR mutations were present. Expression was verified by Western blot with anti-ataxin-3 (Paulson 1997) and GFP antibodies (MBL). The construct encoding a flag tagged, 352 residue tau isoform was previously described (Leger 1994). The pEGFP-tau plasmid was constructed by ligating the human tau cDNA into pEGFP-C2 (Clontech) and encodes tau with EGFP fused to the amino terminus. The pEGFP-tauV337M plasmid was derived using site-directed mutagenesis (QuikChange Kit, Stratagene) of the pEFGP-tau plasmid.

Cell Culture and Transfections. Culture of Cos-7 and HeLa cells has been described (Chai 1999b). Transfections with plasmids and siRNA were performed using Lipofectamine Plus (LifeTechnologies) according to the manufacturer's instructions. For ataxin-3 expression 1.5 μg plasmid was transfected with 5μg *in vitro* synthesized siRNAs. For Tau experiments 1μg plasmid was transfected with 2.5μg siRNA. For expression of hairpin siRNA from the phU6 constructs, 1μg ataxin-3 expression plasmid was transfected with 4μg phU6-siC10i or phU6-siG10i. Cos-7 cells infected with siRNA-expressing adenovirus were transfected with 0.5μg of each expression plasmid.

Stably transfected, doxycycline-inducible cell lines were generated in a subclone of PC12 cells, PC6-3, because of its strong neural differentiation properties (Pittman 19938). A PC6-3 clone stably expressing Tet repressor plasmid (provided by S. Strack, Univ. of Iowa), was transfected with pcDNA5/TO-ataxin-3(Q28) or pcDNA5/TO-ataxin-3(Q166) (Invitrogen). After selection in hygromycin, clones were characterized by Western blot and

immunofluorescence. Two clones, PC6-3-ataxin3(Q28)#33 and PC6-3-ataxin3(Q166)#41, were chosen because of their tightly inducible, robust expression of ataxin-3.

siRNA Plasmid and Viral Production. Plasmids expressing ataxin-3 shRNAs were generated by insertion of head-to-head 21 bp hairpins in phU6 that corresponded to siC10 and siG10 (Xia 2002).

Recombinant adenovirus expressing ataxin-3 specific shRNA were generated from phU6-C10i (encoding C10 hairpin siRNA) and phU6si-G10i (encoding G10 hairpin siRNA) as previously described (Xia 2002, Anderson 2000).

Western Blotting and Immunofluorescence. Cos-7 cells expressing ataxin-3 were harvested 24-48 hours after transfection (Chai 1999b). Stably transfected, inducible cell lines were harvested 72 hours after infection with adenovirus. Lysates were assessed for ataxin-3 expression by Western blot analysis as previously described (Chai 1999b), using polyclonal rabbit anti-ataxin-3 antisera at a 1:15,000 dilution or 1C2 antibody specific for expanded polyQ tracts (Trottier 1995) at a 1:2,500 dilution. Cells expressing Tau were harvested 24 hours after transfection. Protein was detected with an affinity purified polyclonal antibody to a human tau peptide (residues 12-24) at a 1:500 dilution. Anti-alpha-tubulin mouse monoclonal antibody (Sigma St. Louis, MO) was used at a 1:10,000 dilution and GAPDH mouse monoclonal antibody (Sigma St. Louis, MO) was used at a 1:1,000 dilution.

Immunofluorescence for ataxin-3 (Chai 1999b) was carried out using 1C2 antibody (Chemicon International Temecula, CA) at 1:1,000 dilution 48 hours after transfection. Flag-tagged, wild type tau was detected using mouse monoclonal antibody (Sigma St. Louis, MO) at 1:1,000 dilution 24 hours after transfection. Both proteins were detected with rhodamine conjugated secondary antibody at a 1:1,000 dilution.

Fluorescent Imaging and Quantification. Fixed samples were observed with a Zeiss Axioplan fluorescence microscope. Digital images were collected on separate red, green and blue fluorescence channels using a SPOT digital

camera. Images were assembled and overlaid using Adobe Photoshop 6.0. Live cell images were collected with a Kodak MDS 290 digital camera mounted to an Olympus (Tokyo, Japan) CK40 inverted microscope. Fluorescence was quantitated by collecting 3 non-overlapping images per well at low power (10x). Pixel count and intensity for each image was determined using Bioquant Nova Prime software (BIOQUANT Image Analysis Corporation). Background was subtracted by quantitation of images from cells of equivalent density under identical fluorescent illumination. Mock transfected cells were used to assess background fluorescence for all experiments and were stained with appropriate primary and secondary antibodies for simulated heterozygous experiments. Average fluorescence is reported from 2 to 3 independent experiments. The mean of 2 to 3 independent experiments for cells transfected with the indicated expression plasmid and siMiss was set at one. Errors bars depict variation between experiments as standard error of the mean. In simulated heterozygous experiments, a blinded observer scored cells with a positive fluorescence signal for expression of wild type, mutant or both proteins in random fields at high power for two independent experiments. More than 100 cells were scored in each experiment and reported as number of cells with co-expression divided by total number of transfected cells.

Results

Direct Silencing of Expanded Alleles. The inventors first attempted suppression of mutant polyQ expression using siRNA complementary to the CAG repeat and immediately adjacent sequences to determine if the expanded repeat differentially altered the susceptibility of the mutant allele to siRNA inhibition (Figure 6). HeLa cells were transfected with various *in vitro* synthesized siRNAs (Danze 2002) and plasmids encoding normal or expanded polyQ fused to red or green fluorescent protein, respectively (Q19-RFP and Q80-GFP) (Fig. 5a). In negative control cells transfected with Q80-GFP, Q19-RFP and a mistargeted siRNA (siMiss), Q80-GFP formed aggregates (Onodera 1997) which recruited the normally diffuse Q19-RFP (Fig 5a). When the

experiment was performed with siRNA targeted to GFP as a positive control for allele specific silencing, Q80-GFP expression was nearly abolished while Q19-RFP continued to be expressed as a diffusely distributed protein (Fig. 5a). When Q19-RFP and Q80-GFP were co-transfected with siRNA directly targeting the CAG repeat (siCAG) (Fig. 5a) or an immediately adjacent 5' region (data not shown), expression of both proteins was efficiently suppressed.

To test whether siRNA could selectively silence expression of a full-length polyQ disease protein, siRNAs were designed that target the transcript encoding ataxin-3, the disease protein in Machado-Joseph Disease, also known as Spinocerebellar Ataxia Type 3 (MJD/SCA3) (Zoghbi 2000) (Fig. 5b). In transfected cells, siRNA directed against three separate regions -- the CAG repeat, a distant 5' site, or a site just 5' to the CAG repeat (siN'CAG) -- resulted in efficient, but not allele-specific, suppression of ataxin-3 containing normal or expanded repeats (data not shown). Consistent with an earlier study using longer dsRNA (Caplen 2002) the present results show that expanded CAG repeats and adjacent sequences, while accessible to RNAi, may not be preferential targets for silencing.

Allele-specific Silencing of the Mutant PolyQ Gene in MJD/SCA3. In further efforts to selectively inactivate the mutant allele the inventors took advantage of a SNP in the *MJD1* gene, a G to C transition immediately 3' to the CAG repeat (G987C) (Fig. 5b). This SNP is in linkage disequilibrium with the disease-causing expansion, in most families segregating perfectly with the disease allele. Worldwide, 70% of disease chromosomes carry the C variant (Gaspar 2001). The present ataxin-3 expression cassettes, which were generated from patients (Paulson 1997), contain the C variant in all expanded ataxin-3 constructs and the G variant in all normal ataxin-3 constructs. To test whether this G-C mismatch could be distinguished by siRNA, siRNAs were designed that included the last 2 CAG triplets of the repeat followed by the C variant at position 7 (siC7) (Figure 6 and Fig. 5b), resulting in a perfect match only for expanded alleles. Despite the presence of a single mismatch to the wild type allele, siC7 strongly inhibited expression of both alleles (Fig. 5c,d). A second G-

C mismatch was then introduced at position 8 such that the siRNA contained two mismatches as compared to wild type and only one mismatch as compared to mutant alleles (siC7/8). The siC7/8 siRNA effectively suppressed mutant ataxin-3 expression, reducing total fluorescence to an average 8.6% of control levels, with only modest effects on wild type ataxin-3 (average 75.2% of control). siC7/8 also nearly eliminated the accumulation of aggregated mutant ataxin-3, a pathological hallmark of disease (Chan 2000) (Fig. 5d).

To optimize differential suppression, siRNAs were designed containing a more centrally placed mismatch. Because the center of the antisense strand directs cleavage of target mRNA in the RNA Induced Silencing Complex (RISC) complex (Elbashir 2001c), it was reasoned that central mismatches might more efficiently discriminate between wild type and mutant alleles. siRNAs were designed that place the C of the SNP at position 10 (siC10), preceded by the final three triplets in the CAG repeat (Figure 6 and Fig. 5b). In transfected cells, siC10 caused allele-specific suppression of the mutant protein (Fig. 5c,d). Fluorescence from expanded Atx-3-Q166-GFP was dramatically reduced (7.4% of control levels), while fluorescence of Atx-3-Q28-GFP showed minimal change (93.6% of control; Fig. 5c,d). Conversely, siRNA engineered to suppress only the wild type allele (siG10) inhibited wild type expression with little effect on expression of the mutant allele (Fig. 5c,d). Inclusion of three CAG repeats at the 5' end of the siRNA did not inhibit expression of Q19-GFP, Q80-GFP, or full-length ataxin-1-Q30 proteins that are each encoded by CAG repeat containing transcripts (Fig. 7).

In the disease state, normal and mutant alleles are simultaneously expressed. In plants and worms, activation of RNAi against one transcript results in the spread of silencing signals to other targets due to RNA-dependent RNA polymerase (RDRP) activity primed by the introduced RNA (Fire 1998, Tang 2003). Although spreading has not been detected in mammalian cells and RDRP activity is not required for effective siRNA inhibition (Chiu 2002, Schwarz 2002, Martinez 2002), most studies have used cell-free systems in which a mammalian RDRP could have been inactivated. If triggering the mammalian

RNAi pathway against one allele activates cellular mechanisms that also silence the other allele, then siRNA applications might be limited to non-essential genes. To test this possibility, the heterozygous state was simulated by co-transfecting Atx-3-Q28-GFP and Atx-3-Q166 and analyzing suppression by Western blot. As shown in Fig. 5e each siRNA retained the specificity observed in separate transfections: siC7 inhibited both alleles, siG10 inhibited only the wild type allele, and siC7/8 and siC10 inhibited only mutant allele expression.

Effective siRNA therapy for late onset disease will likely require sustained intracellular expression of the siRNA. Accordingly, the present experiments were extended to two intracellular methods of siRNA production and delivery: expression plasmids and recombinant virus (Brummelkamp 2002, Xia 2002). Plasmids were constructed expressing siG10 or siC10 siRNA from the human U6 promoter as a hairpin transcript that is processed intracellularly to produce siRNA (Brummelkamp 2002, Xia 2002). When co-transfected with ataxin-3-GFP expression plasmids, phU6-G10i and phU6-C10i-siRNA plasmids specifically suppressed wild type or mutant ataxin-3 expression, respectively (Fig. 5f).

This result encouraged the inventors to engineer recombinant adenoviral vectors expressing allele-specific siRNA (Xia 2002). Viral-mediated suppression was tested in Cos-7 cells transiently transfected with both Atx-3-Q28-GFP and Atx-3-Q166 to simulate the heterozygous state. Cos-7 cells infected with adenovirus encoding siG10, siC10 or negative control siRNA (Ad-G10i, Ad-C10i, and Ad-LacZi respectively) exhibited allele-specific silencing of wild type ataxin-3 expression with Ad-G10i and of mutant ataxin-3 with Ad-C10i (Fig 8a,b,c). Quantitation of fluorescence (Fig. 8b) showed that Ad-G10i reduced wild type ataxin-3 to 5.4% of control levels while mutant ataxin-3 expression remained unchanged. Conversely, Ad-C10i reduced mutant ataxin-3 fluorescence levels to 8.8% of control and retained 97.4% of wild type signal. These results were confirmed by Western blot where it was further observed that Ad-G10i virus decreased endogenous (primate) ataxin-3 while Ad-C10i did not (Fig 8c).

Viral mediated suppression was also assessed in differentiated PC12 neural cell lines that inducibly express normal (Q28) or expanded (Q166) mutant ataxin-3. Following infection with Ad-G10i, Ad-C10i, or Ad-LacZi, differentiated neural cells were placed in doxycycline for three days to induce maximal expression of ataxin-3. Western blot analysis of cell lysates confirmed that the Ad-G10i virus suppressed only wild type ataxin-3, Ad-C10i virus suppressed only mutant ataxin-3, and Ad-LacZi had no effect on either normal or mutant ataxin-3 expression (Fig. 8d). Thus, siRNA retains its efficacy and selectivity across different modes of production and delivery to achieve allele-specific silencing of ataxin-3.

Allele-Specific Silencing of a Missense Tau Mutation. The preceding results indicate that, for DNA repeat mutations in which the repeat itself does not present an effective target, an associated SNP can be exploited to achieve allele-specific silencing. To test whether siRNA works equally well to silence disease-causing mutations directly, the inventors targeted missense Tau mutations that cause FTDP-17 (Poorkaj 1998, Hutton 1998). A series of 21-24 nt siRNAs were generated *in vitro* against four missense FTDP-17 mutations: G272V, P301L, V337M, and R406W (Figure 6 and Fig 9a). In each case the point mutation was placed centrally, near the likely cleavage site in the RISC complex (position 9, 10 or 11) (Laccone 1999). A fifth siRNA designed to target a 5' sequence in all Tau transcripts was also tested. To screen for siRNA-mediated suppression, the inventors co-transfected GFP fusions of mutant and wild type Tau isoforms together with siRNA into Cos-7 cells. Of the five targeted sites, the inventors obtained robust suppression with siRNA corresponding to V337M (Figure 6 and Fig. 9A) (Poorkaj 1998, Hutton 1998), and thus focused further analysis on this mutation. The V337M mutation is a G to A base change in the first position of the codon (GTG to ATG), and the corresponding V337M siRNA contains the A missense change at position 9 (siA9). This intended V337M-specific siRNA preferentially silenced the mutant allele but also caused significant suppression of wild type Tau (Fig. 9b,c).

Based on the success of this approach with ataxin-3, the inventors designed two additional siRNAs that contained the V337M (G to A) mutation at position 9 as well as a second introduced G-C mismatch immediately 5' to the mutation (siA9/C8) or three nucleotides 3' to the mutation (siA9/C12), such that the siRNA now contained two mismatches to the wild type but only one to the mutant allele. This strategy resulted in further preferential inactivation of the mutant allele. One siRNA, siA9/C12, showed strong selectivity for the mutant tau allele, reducing fluorescence to 12.7% of control levels without detectable loss of wild type Tau (Fig. 9b,c). Next, we simulated the heterozygous state by co-transfecting V337M-GFP and flag-tagged WT-Tau expression plasmids (Fig. 10). In co-transfected HeLa cells, siA9/C12 silenced the mutant allele (16.7% of control levels) with minimal alteration of wild type expression assessed by fluorescence (Fig. 10a) and Western blot (Fig. 10b). In addition, siA9 and siA9/C8 displayed better allele discrimination than we had observed in separate transfections, but continued to suppress both wild type and mutant tau expression (Fig. 10a,b,c).

Discussion

Despite the rapidly growing siRNA literature, questions remain concerning the design and application of siRNA both as a research tool and a therapeutic strategy. The present study, demonstrating allele-specific silencing of dominant disease genes, sheds light on important aspects of both applications.

Because many disease genes encode essential proteins, development of strategies to exclusively inactivate mutant alleles is important for the general application of siRNA to dominant diseases. The present results for two unrelated disease genes demonstrate that in mammalian cells it is possible to silence a single disease allele without activating pathways analogous to those found in plants and worms that result in the spread of silencing signals (Fire 1998, Tang 2003).

In summary, siRNA can be engineered to silence expression of disease alleles differing from wild type alleles by as little as a single nucleotide. This

approach can directly target missense mutations, as in frontotemporal dementia, or associated SNPs, as in MJD/SCA3. The present stepwise strategy for optimizing allele-specific targeting extends the utility of siRNA to a wide range of dominant diseases in which the disease gene normally plays an important or essential role. One such example is the polyglutamine disease, Huntington disease (HD), in which normal HD protein levels are developmentally essential. (Nasir 1995). The availability of mouse models for many dominant disorders, including MJD/SCA3 (Cemal 2002), HD (Lin 2001), and FTDP-17 (Tanemura 2002), allows for the *in vivo* testing of siRNA-based therapy for these and other human diseases.

Example 3

Therapy for DYT1 dystonia: Allele-specific silencing of mutant TorsinA

DYT1 dystonia is the most common cause of primary generalized dystonia. A dominantly inherited disorder, DYT1 usually presents in childhood as focal dystonia that progresses to severe generalized disease. With one possible exception, all cases of DYT1 result from a common GAG deletion in *TOR1A*, eliminating one of two adjacent glutamic acids near the C-terminus of the protein TorsinA (TA). Although the precise cellular function of TA is unknown, it seems clear that mutant TA (TAmut) acts through a dominant-negative or dominant-toxic mechanism. The dominant nature of the genetic defect in DYT1 dystonia suggests that efforts to silence expression of TAmut should have potential therapeutic benefit.

Several characteristics of DYT1 make it an ideal disease in which to explore siRNA-mediated gene silencing as potential therapy. Of greatest importance, the dominant nature of the disease suggests that a reduction in mutant TA, whatever the precise pathogenic mechanism proves to be, will be helpful. Moreover, the existence of a single common mutation that deletes a full three nucleotides suggests it may be feasible to design siRNA that will specifically target the mutant allele and will be applicable to all affected persons.

Finally, there is no effective therapy for DYT1, a relentless and disabling disease. Thus, any therapeutic approach with promise needs to be explored. Because TAwt may be an essential protein, however, it is critically important that efforts be made to silence only the mutant allele.

In the studies reported here , the inventors explored the utility of siRNA for DYT1. As outlined in the strategy in Figure 11, the inventors sought to develop siRNA that would specifically eliminate production of protein from the mutant allele. By exploiting the three base pair difference between wild type and mutant alleles, the inventors successfully silenced expression of TAmut without interfering with expression of the wild type protein (TAwt).

Methods

siRNA design and synthesis Small-interfering RNA duplexes were synthesized *in vitro* according to a previously described protocol (Donze 2002), using AmpliScribeT7 High Yield Transcription Kit (Epicentre Technologies) and desalted DNA oligonucleotides (IDT). siRNAs were designed to target different regions of human TA transcript: 1) an upstream sequence common to both TAwt and TAmut (com-siRNA); 2) the area corresponding to the mutation with either the wild type sequence (wt-siRNA) or the mutant sequence positioned at three different places (mutA-siRNA, mutB-siRNA, mutC-siRNA); and 3) a negative control siRNA containing an irrelevant sequence that does not target any region of TA (mis-siRNA). The design of the primers and targeted sequences are shown schematically in Figure 12. After *in vitro* synthesis, the double stranded structure of the resultant RNA was confirmed in 1.5 % agarose gels and RNA concentration determined with a SmartSpect 3000 UV Spectrophotometer (BioRad).

Plasmids pcDNA3 containing TAwt or TAmut cDNA were kindly provided by Xandra Breakefield (Mass General Hospital, Boston, MA). This construct was produced by cloning the entire coding sequences of human TorsinA (1-332), both wild-type and mutant (GAG deleted), into the mammalian expression vector, pcDNA3 (Clontech, Palo Alto, CA). Using PCR based

strategies, an N-terminal hemagglutinin (HA) epitope tag was inserted into both constructs. pEGFP-C3-TAwt was kindly provided by Pullanipally Shashidharan (Mt Sinai Medical School, NY). This construct was made by inserting the full-length coding sequence of wild-type TorsinA into the EcoRI and BamHI restriction sites of the vector pEGFP-C3 (Clontech). This resulted in a fusion protein including eGFP, three "stuffer" amino acids and the 331 amino acids of TorsinA. HA-tagged TAmut was inserted into the ApaI and SalI restriction sites of pEGFP-C1 vector (Clontech), resulting in a GFP-HA-TAmut construct.

Cell culture and transfections Methods for cell culture of Cos-7 have been described previously (Chai 1999b). Transfections with DNA plasmids and siRNA were performed using Lipofectamine Plus (LifeTechnologies) according to the manufacturer's instructions in six or 12 well plates with cells at 70-90% confluence. For single plasmid transfection, 1 μg of plasmid was transfected with 5μg of siRNA. For double plasmid transfection, 0.75 μg of each plasmid was transfected with 3.75 μg of siRNA.

Western Blotting and Fluorescence Microscopy. Cells were harvested 36 to 48 hours after transfection and lysates were assessed for TA expression by Western Blot analysis (WB) as previously described (Chai 1999b). The antibody used to detect TA was polyclonal rabbit antiserum generated against a TA-maltose binding protein fusion protein (kindly provided by Xandra Breakefield) at a 1:500 dilution. Additional antibodies used in the experiments described here are the anti-HA mouse monoclonal antibody 12CA5 (Roche) at 1:1,000 dilution, monoclonal mouse anti-GFP antibody (MBL) at 1:1,000 dilution, and for loading controls, anti α-tubulin mouse monoclonal antibody (Sigma) at 1:20,000 dilution.

Fluorescence visualization of fixed cells expressing GFP-tagged TA was performed with a Zeiss Axioplan fluorescence microscope. Nuclei were visualized by staining with 5μg/ml DAPI at room temperature for 10 minutes. Digital images were collected on separate red, green and blue fluorescence channels using a Diagnostics SPOT digital camera. Live cell images were collected with a Kodak MDS 290 digital camera mounted on an Olympus CK40

inverted microscope equipped for GFP fluorescence and phase contrast microscopy. Digitized images were assembled using Adobe Photoshop 6.0.

Western Blot and Fluorescence Quantification. For quantification of WB signal, blots were scanned with a Hewlett Packard ScanJet 5100C scanner. The pixel count and intensity of bands corresponding to TA and α-tubulin were measured and the background signal subtracted using Scion Image software (Scion Corporation). Using the α-tubulin signal from control lanes as an internal reference, the TA signals were normalized based on the amount of protein loaded per lane and the result was expressed as percentage of TA signal in the control lane. Fluorescence quantification was determined by collecting three non-overlapping images per well at low power (10x), and assessing the pixel count and intensity for each image with Bioquant Nova Prime software (BIOQUANT Image Analysis Corporation). Background fluorescence, which was subtracted from experimental images, was determined by quantification of fluorescence images of untransfected cells at equivalent confluence, taken under identical illumination and exposure settings.

RESULTS

Expression of tagged TorsinA constructs. To test whether allele-specific silencing could be applied to DYT1, a way to differentiate TAwt and TAmut proteins needed to be developed. Because TAwt and TAmut display identical mobility on gels and no isoform-specific antibodies are available, amino-terminal epitope-tagged TA constructs and GFP-TA fusion proteins were generated that would allow distinguishingTAwt and TAmut. The use of GFP-TA fusion proteins also facilitated the ability to screen siRNA suppression because it allowed visualization of TA levels in living cells over time.

In transfected Cos-7 cells, epitope-tagged TA and GFP-TA fusion protein expression was confirmed by using the appropriate anti-epitope and anti-TA antibodies. Fluorescence microscopy in living cells showed that GFP-TAwt and GFP-TAmut fusion proteins were expressed diffusely in the cell, primarily in the cytoplasm, although perinuclear inclusions were also seen. It is important to note

that these construct were designed to express reporter proteins in order to assess allele-specific RNA interference rather than to study TA function. The N-terminal epitope and GFP domains likely disrupt the normal signal peptide-mediated translocation of TA into the lumen of the endoplasmic reticulum, where TA is thought to function. Thus, while these constructs facilitated expression analysis in the studies described here, they are of limited utility for studying TA function.

Silencing TorsinA with siRNA. Various siRNAs were designed to test the hypothesis that siRNA-mediated suppression of TA expression could be achieved in an allele-specific manner (figure 12). Because siRNA can display exquisite sequence specificity, the three base pair difference between mutant and wild type TOR1A alleles might be sufficient to permit the design of siRNA that preferentially recognizes mRNA derived from the mutant allele. Two siRNAs were initially designed to target TAmut (mutA-siRNA and mutB-siRNA) and one to target TAwt (wt-siRNA). In addition, a positive control siRNA was designed to silence both alleles (com-siRNA) and a negative control siRNA of irrelevant sequence (mis-siRNA) was designed. Cos-7 cells were first cotransfected with siRNA and plasmids encoding either GFP-TAwt or untagged TAwt at a siRNA to plasmid ratio of 5:1. With wt-siRNA, potent silencing of TAwt expression was observed to less than 1 % of control levels, based on western blot analysis of cell lysates (Figures 13A and 13C). With com-siRNA, TAwt expression was suppressed to ~30 % of control levels. In contrast, mutA-siRNA did not suppress TAwt and mutB-siRNA suppressed TAwt expression only modestly. These results demonstrate robust suppression of TAwt expression by wild type-specific siRNA but not mutant-specific siRNA.

To assess suppression of TAmut, the same siRNAs were cotransfected with plasmids encoding untagged or HA-tagged TAmut. With mutA-siRNA or mutB-siRNA, marked, though somewhat variable, suppression of TAmut expression was observed as assessed by western blot analysis of protein levels (Figure 13B and 13C). With com-siRNA, suppression of TAmut expression was observed similar to what was observed with TAwt expression. In contrast, wt-

siRNA did not suppress expression of TAmut. Thus differential suppression of TAmut expression was observed by allele-specific siRNA in precisely the manner anticipated by the inventors.

To achieve even more robust silencing of TAmut, a third siRNA was engineered to target TAmut (mutC-siRNA, Figure 12). MutC-siRNA places the GAG deletion more centrally in the siRNA duplex. Because the central portion of the antisense strand of siRNA guides mRNA cleavage, it was reasoned that placing the GAG deletion more centrally might enhance specific suppression of TAmut. As shown in Figure 13, mutC-siRNA suppressed TAmut expression more specifically and robustly than the other mut-siRNAs tested. In transfected cells, mutC-siRNA suppressed TAmut to less than 0.5% of control levels, and had no effect on the expression of TAwt.

To confirm allele-specific suppression by wt-siRNA and mutC-siRNA, respectively, the inventors cotransfected cells with GFP-TAwt or GFP-TAmut together with mis-siRNA, wt-siRNA or mutC-siRNA. Levels of TA expression were assessed 24 and 48 hours later by GFP fluorescence, and quantified the fluorescence signal from multiple images was quantified. The results (Figure 13D and 13E) confirmed the earlier western blots results in showing potent, specific silencing of TAwt and TAmut by wt-siRNA and mutC-siRNA, respectively, in cultured mammalian cells.

Allele-specific silencing in simulated heterozygous state. In DYT1, both the mutant and wild type alleles are expressed. Once the efficacy of siRNA silencing was established, the inventors sought to confirm siRNA specificity for the targeted allele in cells that mimic the heterozygous state of DYT1. In plants and *Caenorhabditis elegans*, RNA-dependent RNA polymerase activity primed by introduction of exogenous RNA can result in the spread of silencing signals along the entire length of the targeted mRNA (Fire 1998, Tang 2003). No evidence for such a mechanism has been discovered in mammalian cells (Schwarz 2002, Chiu 2002). Nonetheless it remained possible that silencing of the mutant allele might activate cellular processes that would also inhibit expression from the wild type allele. To address this possibility, Cos-7 cells were

cotransfected with both GFP-TAwt and HA-TAmut, and suppression by mis-siRNA, wt-siRNA or mutC-siRNA was assessed. As shown in Figure 14, potent and specific silencing of the targeted allele (either TAmut or TAwt) to levels less than 1% of controls was observed, with only slight suppression in the levels of the non-targeted protein. Thus, in cells expressing mutant and wild type forms of the protein, siRNA can suppress TAmut while sparing expression of TAwt.

DISCUSSION

In this study the inventors succeeded in generating siRNA that specifically and robustly suppresses mutant TA, the defective protein responsible for the most common form of primary generalized dystonia. The results have several implications for the treatment of DYT1 dystonia. First and foremost, the suppression achieved was remarkably allele-specific, even in cells simulating the heterozygous state. In other words, efficient suppression of mutant TA occurred without significant reduction in wild type TA. Homozygous TA knockout mice die shortly after birth, while the heterozygous mice are normal (Goodchild 2002), suggesting an essential function for TA. Thus, therapy for DYT1 needs to eliminate the dominant negative or dominant toxic properties of the mutant protein while sustaining expression of the normal allele in order to prevent the deleterious consequences of loss of TA function. Selective siRNA-mediated suppression of the mutant allele fulfills these criteria without requiring detailed knowledge of the pathogenic mechanism.

An appealing feature of the present siRNA therapy is applicable to all individuals afflicted with DYT1. Except for one unusual case (Leung 2001, Doheny 2002, Klein 2002b), all persons with DYT1 have the same (GAG) deletion mutation (Ozelius 1997, Ozelius 1999). This obviates the need to design individually tailored siRNAs. In addition, the fact that the DYT1 mutation results in a full three base pair difference from the wild type allele suggests that siRNA easily distinguishes mRNA derived from normal and mutant *TOR1A* alleles.

It is important to recognize that DYT1 is not a fully penetrant disease (Fahn 1998, Klein 2002a). Even when expressed maximally, mutant TA causes significant neurological dysfunction less than 50% of the time. Thus, even partial reduction of mutant TA levels might be sufficient to lower its pathological brain activity below a clinically detectable threshold. In addition, the DYT1 mutation almost always manifests before age 25, suggesting that TAmut expression during a critical developmental window is required for symptom onset. This raises the possibility that suppressing TAmut expression during development might be sufficient to prevent symptoms throughout life. Finally, unlike many other inherited movement disorders DYT1 is not characterized by progressive neurodegeneration. The clinical phenotype must result primarily from neuronal dysfunction rather than neuronal cell death (Hornykiewicz 1986, Walker 2002, Augood 2002, Augood 1999). This suggests the potential reversibility of DYT1 by suppressing TAmut expression in overtly symptomatic persons.

Example 4

siRNA Specific for Huntington's Disease

The present inventors have developed huntingtin siRNA focused on two targets. One is non-allele specific (siHDexon2), the other is targeted to the exon 58 codon deletion, the only known common intragenic polymorphism in linkage disequilibrium with the disease mutation (Ambrose et al, 1994). Specifically, 92% of wild type huntingtin alleles have four GAGs in exon 58, while 38% of HD patients have 3 GAGs in exon 58. To assess a siRNA targeted to the intragenic polymorphism, PC6-3 cells were transfected with a full-length huntingtin containing the exon 58 deletion. Specifically, PC6-3 rat pheochromocytoma cells were co-transfected with CMV-human Htt (37Qs) and U6 siRNA hairpin plasmids. Cell extracts were harvested 24 hours later and western blots were performed using 15 μg total protein extract. Primary antibody was an anti-huntingtin monoclonal antibody (MAB2166, Chemicon) that reacts with human, monkey, rat and mouse Htt proteins.

As seen in Figure 15, the siRNA lead to silencing of the disease allele. As a positive control, a non-allele specific siRNA targeted to exon 2 of the huntingtin gene was used. siRNA directed against GFP was used as a negative control. Note that only siEx58# 2 is functional. The sequence for siEX58#2 is the following: 5'-AAGAGGAGGAGGCCGACGCCC-3' (SEQ ID NO:90). siEX58#1 was only minimally functional.

Example 5

RNA Interference Improves Motor and Neuropathological Abnormalities in a Huntington's Disease Mouse Model

Huntington's disease (HD) is one of nine dominant neurodegenerative diseases resulting from polyglutamine repeat expansions (CAG codon, Q) in exon 1 of HD, leading to a toxic gain of function on the protein huntingtin (htt) (The Huntington's Disease Collaborative Research Group (1993) Cell 72, 971-83; Gusella *et al.*, (2000) Nat Rev Neurosci 1, 109-15). Hallmark HD characteristics include cognitive and behavioral disturbance, involuntary movements (chorea), neuronal inclusions, and striatal and cortical neurodegeneration (Gusella *et al.*, (2000) Nat Rev Neurosci 1, 109-15). Htt alleles containing greater than 35 CAG repeats generally cause HD, with age-at-onset correlating inversely with expansion length, a common characteristic of the polyglutamine repeat disorders. The disease usually develops in mid-life, but juvenile-onset cases can occur with CAG repeat lengths greater than 60. Death typically occurs 10-15 years after symptom onset. Currently, no preventative treatment exists for HD.

Therapies aimed at delaying disease progression have been tested in HD animal models. For example, beneficial effects have been reported in animals treated with substances that increase transcription of neuroprotective genes (histone deacetylase) (Ferrante *et al.*, (2003) J Neurosci 23, 9418-27); prevent apoptosis (caspase inhibitors)(Ona *et al.*, (1999) Nature 399, 263-7); enhance energy metabolism (coenzyme Q/remacemide, creatine) (Ferrante *et al.*, (2002) J Neurosci 22, 1592-9; Andreassen *et al.*, (2001) Neurobiol Dis 8, 479-91); and

inhibit the formation of polyglutamine aggregates (trehalose, Congo red, cystamine) (Tanaka *et al.*, (2004) Nat Med 10, 148-54; Karpuj *et al.*, (2002) Nat Med 8, 143-9; Sanchez *et al.*, (2003) Nature 421, 373-9). These approaches target downstream and possibly indirect effects of disease allele expression. In contrast, no therapies have been described that directly reduce mutant huntingtin gene expression, thereby targeting the fundamental, underlying pathological insult.

The therapeutic promise of silencing mutant htt expression was demonstrated in a tetracycline-regulated mouse model of HD (Yamamoto *et al.*, (2000) Cell 101, 57-66). When mutant htt was inducibly expressed, pathological and behavioral features of the disease developed, including the characteristic neuronal inclusions and abnormal motor behavior. Upon repression of transgene expression in affected mice, pathological and behavioral features resolved. Thus, reduction of htt expression using RNAi may allow protein clearance mechanisms within neurons to normalize mutant htt-induced changes. We hypothesize that directly inhibiting the expression of mutant htt will slow or prevent HD-associated symptom onset in a relevant animal model.

Screening of putative therapies for HD has benefited from the existence of several HD mouse models (Beal *et al.*, (2004) Nat Rev Neurosci 5, 373-84; Levine *et al.*, (2004) Trends Neurosci 27, 691-7). HD-like phenotypes are displayed in knock-in mice (Lin *et al.*, (2001) Hum Mol Genet 10, 137-44; Menalled *et al.*, (2003) J Comp Neurol 465, 11-26), drug-induced models (McBride *et al.*, (2004) J Comp Neurol 475, 211-9) and transgenic mice expressing full-length mutant huntingtin (*e.g.* YAC-transgenic mice) (Hodgson *et al.*, (1999) Neuron 23, 181-92; Slow *et al.*, (2003) Hum Mol Genet 12, 1555-67; Reddy *et al.*, (1998) Nat Genet 20, 198-202) or an N-terminal fragment of htt (Yamamoto *et al.*, (2000) Cell 101, 57-66; Mangiarini *et al.*, (1996) Cell 87(3), 493-506; Schilling *et al.*, (1999) Hum Mol Genet 8(3), 397-407). Mice expressing truncated N-terminal fragments of huntingtin have been valuable for proof-of-principle evaluation of therapies because they show rapidly progressive motor abnormalities and striatal neuropathology, phenotypes which do not

develop or develop very late in knock-in or YAC transgenic mice. Mice expressing truncated forms of huntingtin thus replicate more severe forms of the disease. The present inventors tested if RNA interference (RNAi) induced by short hairpin RNAs (shRNAs) (Dykxhoorn *et al.*, (2003) Nat Rev Mol Cell Biol 4, 457-67) could reduce expression of mutant htt and improve HD-associated abnormalities in a transgenic mouse model of HD. It was found that RNAi directed against mutant human huntingtin (htt) reduced htt mRNA and protein expression in cell culture and in HD mouse brain. It is important to note that htt gene silencing improved behavioral and neuropathological abnormalities associated with HD.

Materials and Methods

Plasmids and Adeno-Associated Virus (AAV) construction. Myc-tagged HD-N171-82Q was expressed from a pCMV-HD-N171-82Q plasmid (Schilling *et al.*, (1999) Hum Mol Genet 8(3), 397-407). PCR (Pfu polymerase, Stratagene) was used to amplify the U6 promoter along with shRNAs targeting human huntingtin (shHD2.1; Fig. 16A), eGFP (shGFP) (Xia *et al.*, (2002) Nat Biotechnol 20, 1006-1010); or *E. coli* β-galactosidase (bp 1152-1172; shLacZ). PCR products were cloned, verified by sequencing and inserted into pAAV.CMV.hrGFP, which contains AAV-2 ITRs, a CMV-hrGFP-SV40 polyA reporter cassette, and sequences used for homologous recombination into baculovirus (Urabe *et al.*, (2002) Hum Gene Ther 13, 1935-1943). Recombinant AAV serotype 1 capsid vectors were generated as described (Urabe *et al.*, (2002) Hum Gene Ther 13, 1935-1943). AAV titers were determined by quantitative PCR and/or DNA slot blot and were 5×10^{12} vector genomes/ml.

Animals. All animal studies were approved by the University of Iowa Animal Care and Use Committee. HD-N171-82Q mice were purchased from Jackson Laboratories, Inc. (Schilling *et al.*, (1999) Hum Mol Genet 8(3), 397-407; Schilling *et al.*, (2001) Neurobiol Dis 8, 405-18) and maintained on a

B6C3F1/J background. Heterozygous and age-matched wildtype littermates were used for the experiments, as indicated.

Northern blots. HEK293 cells were transfected (Lipofectamine-2000; Invitrogen) with pCMV-HD-N171-82Q and plasmids expressing shHD2.1, shGFP, or shLacZ at shRNA:target ratios of 8:1. Forty-eight hours post-transfection, RNA was harvested (Trizol Reagent; Invitrogen) and 10 μg were assessed northern blot (NorthernMax; Ambion) using probes to human htt or human GAPDH. Band intensities were quantified using a phosphorimager (Storm 860 instrument and ImageQuant v1.2 software, Molecular Dynamics).

For *in vivo* studies, total RNA was isolated from hrGFP-positive striata. Thirty μg RNA was run on 15% polyacrylamide-urea gels, transferred to Hybond N+ membranes (Amersham Pharmacia), then probed with ^{32}P-labeled sense oligonucleotides at 36°C for 3 h, washed in 2X SSC (36°C), and exposed to film.

Western blots. HEK293 cells were transfected as described with shHD2.1 or shGFP singly or in combination with pCMV-HD-N171-82Q. Forty-eight hours later, cells were lysed to recover total protein. Western blots were incubated with anti-myc (1:5,000; Invitrogen), anti full-length human htt (1:5,000; MAB2166; Chemicon), or anti-human β-actin (1:5,000; Clone AC-15; Sigma) followed by HRP-coupled goat anti-mouse or goat anti-rabbit secondary antibodies (1:20,000 and 1:100,000, respectively; Jackson Immunochemicals). Blots were developed using ECL-Plus reagents (Amersham Biosciences). For evaluation of transduced brain, 3 week old mice were injected as described and protein was harvested from striata 2 weeks later. Twenty-five μg were run on SDS-PAGE gels as described, transferred to nitrocellulose, then probed with antibodies to detect human htt (1:500, mEM48; Gift from X.J. Li) and mouse prion protein (1:40,000; Chemicon International). Secondary antibody incubations were performed as described above.

Quantitative RT-PCR

In vitro shRNA dose response. HEK293 cells were transfected with 0 (mock), 10, 100, or 1000 ng of shLacZ or shHD2.1 and RNA was harvested 24 h later. Following DNase treatment (DNA-Free, Ambion), random-primed, first strand cDNA was generated from 500 ng total RNA (Taqman™ Reverse Transcription Reagents, Applied Biosystems) according to manufacturer's protocol. Taqman™ Assays were performed on an ABI Prism 7000 Sequence Detection System using Taqman™ 2X Universal PCR Master Mix (Applied Biosystems) and Taqman™ primers/probe sets specific for human htt and mammalian rRNA (Applied Biosystems). Relative gene expression was determined using the relative standard curve method.

In vivo huntingtin mRNA expression. Striata were dissected from 5.5 month old mice, snap frozen in liquid nitrogen, and pulverized. cDNA was generated as described above. Relative gene expression was assayed using Taqman™ primers/probe sets specific for human htt and mammalian rRNA or Assays-By-Design Taqman™ primers/probes specific for mouse huntingtin (mHdh; Applied Biosystems). All values were calibrated to contralateral, uninjected striata. For human huntingtin detection; shHD2.1 samples, n=8 striata; shLacZ, n=7; uninjected, n=4. For mouse Hdh detection; injected HD samples, n=4; uninjected samples n=2.

AAV Injections

All animal procedures were pre-approved by the University of Iowa Animal Care and Use Committee. AAV Injections were performed in 4 week old mice using the following parameters (coordinates are reported with respect to the bregma): Striatal: 0.5 mm anterior, 2.5 mm lateral, 2.5 mm depth, 5 µl/site, 250 nl/min infusion rate. Cerebellar: 0.1 mm depth, 1 µl/site, 250 nl/min infusion rate.

Behavioral analysis

Stride length measurements. Mice injected bilaterally at 4 weeks of age were analyzed at 4 months of age. Analyses were performed as described previously (Carter *et al.*, (1999) J Neurosci 19, 3248) with some modifications. Specifically, mice were allowed to walk across a paper-lined chamber measuring 100 cm long, 10 cm wide, with 10 cm high walls into an enclosed box. Mice were given one practice run and were then tested three times to produce three separate footprint tracings, totaling 42 measurements each for front and rear footprints per mouse. Measurements were averaged and data presented as box plots. ANOVA with Scheffe's post-hoc test was performed to determine statistical significance. Uninjected mice, n=4; injected WT, n=3; injected N171-82Q, n=6 mice.

Rotarod performance test. Two separate experimental cohorts of mice were injected at 4 weeks of age and tested on the rotarod (Model 7650, Ugo Basile Biological Research Apparatus) at 10 and 18 weeks of age as previously described (Xia *et al.*, (2004) Nat Med 10, 816-820). Data from trials 2-4 for each day are presented as means ± S.E.M. Uninjected WT, n=6; shLacZ WT, n=5, shHD2.1 WT, n=6; uninjected N171-82Q, n=5; shLacZ N171-82Q, n=10; shHD2.1 N171-82Q, n=11). Reported values are means ± S.E.M.

Immunofluorescence

Forty μm free-floating coronal sections were stained with mEM48 antibody (1:500; 24 h, 4°C), followed by Alexa-568 labeled goat anti-mouse secondary antibody (1:200; 4 h, room temp; Molecular Probes). Sections were mounted onto slides, covered in Gel/Mount (Biomeda Corp) and images were captured using fluorescent microscopy (Leica DM RBE or Zeiss confocal) equipped with a CCD-camera (SPOT RT, Diagnostics Instruments). **Results**

shHD2.1 Reduces Human Huntingtin Expression *In Vitro*

In vitro screening was used to identify effective shRNAs directed against a CMV-promoter transcribed HD-N171-82Q mRNA, which is identical to the pathogenic truncated huntingtin fragment transgene present in HD-N171-82Q mice (Schilling *et al.*, (1999) Hum Mol Genet 8(3), 397-407). Hairpin

constructs targeting sequences in human exons 1-3 were evaluated by co-transfection. One htt-targeted shRNA, shHD2.1 (Fig. 16*A*), reduced HD-N171-82Q mRNA and protein levels by ~85 and ~55% respectively, relative to control shRNA treated samples (Fig. 16*B, C*). Interestingly, none of the shRNAs tested that targeted exon 1 were functional under these conditions and in this system. Additional siRNAs can be screened as described herein to identify functional siRNAs targeting exon 1 of the HD gene in this other other systems.

To test if shHD2.1 could silence endogenous full-length human htt expression, HEK 293 cells were transfected with plasmids expressing shHD2.1 or shGFP. ShHD2.1, but not control shRNAs, directed gene silencing of endogenous htt mRNA and protein (Figs. 16*D, E*). This system can be readily used to screen additional siRNAs targeting the HD gene.

Expression of shRNA in Mouse Brain

Next, the inventors tested U6 promoter-transcribed shHD2.1 expression *in vivo* and determined its effects on HD-associated symptoms in mice. This pol III dependent promoter has not previously been evaluated in striata for sustained expression *in vivo*, although shRNAs have been expressed in brain using either the pol II-dependent CMV promoter in striatum (Xia *et al.*, (2002) Nat Biotechnol 20, 1006-1010) or the H1 promoter in cerebellar degeneration models (Xia *et al.*, (2004) Nat Med 10, 816-820). U6 promoter-driven shHD2.1, and the control hairpin shLacZ, were cloned into adeno-associated virus (AAV) shuttle plasmids that contained a separate CMV-humanized *Renilla* green fluorescent protein (hrGFP) reporter cassette (Fig. 17*A*). High-titer AAV1 particles (AAV.shHD2.1 and AAV.shLacZ), which have broad neuronal tropism, were generated (Urabe *et al.*, (2002) Hum Gene Ther 13, 1935-1943), and hairpin expression was assessed after injection into mouse striatum. The N171-82Q mouse model was used because shHD2.1 targets sequences in exon 2, precluding use of the R6/2 transgenic model, which expresses only exon 1 of the HD gene. As shown in Fig. 17*B*, precursor and processed shRNAs (~50 nt and 21 nt, respectively) were expressed three weeks after transduction, indicating sustained

expression and appropriate processing of shRNAs in the striatum. Analysis of coronal brain sections from injected mice showed widespread transduction (Fig. 17*C*; hrGFP fluorescence) up to 5 months post-injection.

AAV.shHD2.1 Reduces HD-N171-82Q Expression *In Vivo*

The inventors next investigated the effects of RNAi on the characteristic HD-associated neuronal inclusions and HD-N171-82Q mRNA levels *in vivo*. Tissues were harvested from end-stage HD-N171-82Q mice (~5.5 months of age) because striatal inclusions are less robust at earlier ages in this model. In striata from HD-N171-82Q mice injected with AAV.shHD2.1, htt-reactive inclusions were absent in transduced cells compared to untransduced regions (Fig. 18*A*, lower panels; Fig. 18*B*). Conversely, abundant inclusions were detected in transduced regions from AAV.shLacZ-injected HD mice (Fig. 18*A*, upper panels). No inclusions were observed in WT mice (data not shown). In addition, western analysis revealed that soluble HD-N171-82Q monomer was decreased in mouse striata transduced with AAV.shHD2.1 compared to uninjected or AAV.shLacZ-injected controls (Fig. 18*C*). The reduction in protein levels detected by immunohistochemistry and western blot was due to decreased transgene expression. HD-N171-82Q mRNA was reduced 51% to 55% in AAV.shHD2.1-injected HD mice relative to AAV.shLacZ-injected or uninjected HD mice (Fig. 18*D*). AAV.shHD2.1 and AAV.shLacZ had no effect on endogenous mouse htt expression (Avg. mHDH expression: Uninjected HD, 1.00±0.09; Uninjected WT, 1.13±0.04; AAV.shLacZ injected HD, 1.10±0.08; AAV.shHD2.1 injected HD, 1.08±0.05).

Neuronal inclusions in HD-N171-82Q striata are variable. Inclusions may be present in as few as 10% and up to 50% of all striatal neurons in different end-stage HD-N171-82Q mice (Schilling *et al.*, (1999) Hum Mol Genet 8(3), 397-407). In contrast, robust and widespread EM48-positive inclusions are present in cerebellar granule cells by ~3 months of age [(Schilling *et al.*, (1999) Hum Mol Genet 8(3), 397-407) and Fig. 18], and cerebellar HD-N171-82Q mRNA levels are ~8 fold higher relative to striatum (QPCR, data not shown).

This high-level cerebellar expression is partially attributable to the transcriptional profile of the prion promoter driving HD-N171-82Q transgene expression (Schilling *et al.*, (1999) Hum Mol Genet 8(3), 397-407). Cerebellar inclusions are not typically found in brains of adult-onset HD patients. However, cerebellar pathology has been reported in juvenile onset HD cases, which are the most severe forms of the disease, and interestingly, in Hdh140 knock-in mice as early as 4 months of age (Menalled *et al.*, (2003) J Comp Neurol 465, 11-26; Nance *et al.*, (2001) Ment Retard Dev Disabil Res Rev 7, 153-7; Fennema- *et al.*, (2004) Neurology 63, 989-95; Seneca *et al.*, (2004) Eur J Pediatr.; Byers *et al.*, (1973) Neurology 23, 561-9; Wheeler *et al.*, (2002) Hum Mol Genet 11, 633-40). The abundant inclusions in HD-N171-82Q cerebellar neurons provide a second target for assessing the effects of AAV.shHD2.1 on target protein levels. Direct cerebellar injections were done into a separate cohort of mice, and HD-N171-82Q expression examined by immunofluorescence. Together the data show that AAV.shHD2.1, but not control AAV.shLacZ, reduces mutant htt expression and prevents formation of the disease-associated neuronal inclusions.

Striatal Delivery of AAV.shHD2.1 improves established behavioral phenotypes

The effects of shRNA treatment on established behavioral deficits and animal weight were tested. RNAi directed to striatum did not normalize the notable weight differences between HD-N171-82Q and WT mice (shHD2.1-injected, 22.7±3.8 g; shLacZ, 22.6±2.8 g; compared to age-matched wild-type mice (shHD2.1, 26.3±0.4; shLacZ, 27.3±5.8), confirming that intracerebral injection confines RNAi therapy to the site of application (Schilling *et al.*, (1999) Hum Mol Genet 8(3), 397-407; Xia *et al.*, (2004) Nat Med 10, 816-820). However, significant improvements in stride length measurements and rotarod deficits were noted.

Stride length and rotarod tests were performed on uninjected mice, and mice injected bilaterally into striatum with AAVshHD2.1 or AAVshLacZ. As

shown in Fig. 19*A*, HD-N171-82Q mice display significantly shorter stride lengths than those of wild-type (WT) mice, consistent with prior work (Menalled *et al.*, (2003) J Comp Neurol 465, 11-26; Carter *et al.*, (1999) J Neurosci 19, 3248; Wheeler *et al.*, (2002) Hum Mol Genet 11, 633-40). Gait deficits in AAV.shHD2.1-treated HD-N171-82Q mice were significantly improved compared to AAV.shLacZ-treated (improvements for front and rear strides, 13 and 15%, respectively; $p<0.0001$) and uninjected HD-N171-82Q mice (front and rear strides, 14 and 18%, respectively; $p<0.0001$). Gait improvements did not fully resolve, as all HD-N171-82Q groups remained significantly different than their age-matched WT littermates. There was no effect of AAV.shLacZ or AAV.shHD2.1 expression on stride lengths of WT mice.

The accelerating rotarod test was used to confirm the beneficial behavioral effects of RNAi targeted to the mutant human HD allele (Schilling *et al.*, (1999) Hum Mol Genet 8(3), 397-407). Mice were left uninjected, or were injected bilaterally into the striatum with AAV.shLacZ or AAV.shHD2.1 at 4 weeks of age, followed by rotarod analyses at 10- and 18-weeks of age (Fig. 19*B*). By 10 weeks, uninjected and AAV.shLacZ-injected HD mice show impaired performance relative to all other groups, and continued to demonstrate significantly reduced performance over the course of the study ($p<0.05$ relative to all other groups). It is important to note that HD mice treated with AAVshHD2.1 showed dramatic behavioral improvements relative to control-treated HD mice ($p<0.0008$) (Fig. 19*B*). AAV.shLacZ-treated HD mice showed a 22% decline ($p<0.005$; ANOVA), while AAV.shHD2.1-treated HD mice displayed a modest, non-significant 3% drop in rotarod performance between 10 and 18 weeks of age. There was a partial normalization of rotarod deficits in HD mice injected with AAV.shHD2.1 compared to WT mice that was consistent with the gait analyses.

The inventors found no decline in stride length or rotarod performance between WT mice left untreated, or those injected with shRNA-expressing AAVs (Fig. 19*A*,*B*). However, at 10 weeks, there was a dramatic difference in rotarod performance between uninjected WT and all groups of injected WT

mice, which resolved by 18 weeks of age. These data suggest that there was some detrimental effect of direct brain injection on rotarod performance from which the mice recovered over time. These data suggest that RNAi expression in mammalian brain had no overt negative impact on motor behavior (Fig. 19*A,B*).

Discussion

The inventors have shown that motor and neuropathological abnormalities in a relevant HD mouse model are significantly improved by reducing striatal expression of a pathogenic huntingtin allele using AAV1-delivered shRNA. The inventors have previously shown that RNAi can improve neuropathology and behavioral deficits in a mouse model of spino-cerebellar ataxia type 1 (SCA1) (Xia *et al.*, (2004) Nat Med 10, 816-820), a dominant neurodegenerative disorder that affects a population of neurons distinct from those degenerating in HD.

The shHD2.1 hairpin sequence reduced huntingtin expression *in vitro* and *in vivo*, and it is important to note, the present northern blot data suggest that the processed active guide strand was protected by RISC *in vivo*. The activity of the shRNAs could be improved using recently described rules for optimal shRNA design (Reynolds *et al.*, (2004) Nat Biotechnol 22, 326-30; Schwarz *et al.*, (2003) Cell 115, 199-208; Khvorova *et al.*, (2003) Cell 115, 505; Ui-Tei *et al.*, (2004) Nucleic Acids Res 32, 936-48).

Prior work demonstrated an essential role for huntingtin in embryogenesis and postnatal neurogenesis (Nasir *et al.*, (1995) Cell 81, 811-23; Duyao *et al.*, (1995) Science 269, 407-10; White *et al.*, (1997) Nat Genet 17, 404-10; Dragatsis *et al.*, (2000) Nat Genet 26, 300-6). However the effect of partial reduction of normal huntingtin expression in adult, post-mitotic neurons *in vivo* is unknown. In the current study, shHD2.1 reduced expression of a mutant, disease-causing human htt transgene, but had no effect on normal mouse huntingtin expression due to sequence differences between mouse and human genes. In HD patients, shHD2.1 would be expected to reduce expression of both

the mutant and normal huntingtin alleles. The present data show that HD-like symptoms can be improved by even a partial reduction of mutant htt expression, suggesting that complete elimination of mutant allele expression may not be required.

In summary, the inventors have shown that RNAi can dramatically improve HD-associated abnormalities, including pathological and behavioral deficits, in a HD mouse model.

Example 6

Huntington's Disease (HD)

Huntington's disease (HD) is one of several dominant neurodegenerative diseases that result from a similar toxic gain of function mutation in the disease protein: expansion of a polyglutamine (polyQ)-encoding tract. It is well established that for HD and other polyglutamine diseases, the length of the expansion correlates inversely with age of disease onset. Animal models for HD have provided important clues as to how mutant huntingtin (htt) induces pathogenesis. Currently, no neuroprotective treatment exists for HD. RNA interference has emerged as a leading candidate approach to reduce expression of disease genes by targeting the encoding mRNA for degradation.

Short hairpin RNAs (shRNAs) were generated that significantly inhibited human htt expression in cell lines. Importantly, the shRNAs were designed to target sequences present in HD transgenic mouse models. The present studies test the efficacy of the shRNAs in HD mouse models by determining if inclusions and other pathological and behavioral characteristics that are representative of HD can be inhibited or reversed. In a transgenic model of inducible HD, pathology and behavior improved when mutant gene expression was turned off. These experiments show that RNAi can prevent or reverse disease.

Although the effect of partial reduction of wildtype htt in adult neurons is unknown, it is advantageous to target only mutant htt for degradation, if possible. One polymorphism in linkage disequilibrium with HD has been

identified in the coding sequence for htt, and others are currently being investigated. Disease allele-specific RNAi are designed using approaches that led to allele specific silencing for other neurogenetic disease models. This would allow directed silencing of the mutant, disease-causing expanded allele, leaving the normal allele intact.

Constitutive expression of shRNA can prevent the neuropathological and behavioral phenotypes in a mouse model of Spinocerebellar Ataxia type I, a related polyQ disease. However, the constitutive expression of shRNA may not be necessary, particularly for pathologies that take many years to develop but may be cleared in a few weeks or months. For this reason, and to reduce long-term effects that may arise if nonspecific silencing or activation of interferon responses is noted, controlled expression may be very important. In order to regulate RNAi for disease application, doxycycline-responsive vectors have been developed for controlled silencing in vitro.

HD researchers benefit from a wealth of animal models including six transgenic and four knock-in mouse models (Bates 2003). Expression is from the endogenous human promoter, and the CAG expansion in the R6 lines ranges from 110 to approximately 150 CAGs. The R6/2 line is the most extensively studied line from this work. R6/2 mice show aggressive degenerative disease, with age of symptom onset at 8-12 weeks, and death occurring at 10 to 13 weeks. Neuronal intranuclear inclusions, a hallmark of HD patient brain, appear in the striatum and cortex of the R6/2 mouse (Meade 2002).

Adding two additional exons to the transgene and restricting expression via the prion promoter led to an HD mouse model displaying important HD characteristics but with less aggressive disease progression (Shilling 1999, Shilling 2001). The Borchelt model, N171-82Q, has greater than wildtype levels of RNA, but reduced amounts of mutant protein relative to endogenous htt. N171-82Q mice show normal development for the first 1-2 months, followed by failure to gain weight, progressive incoordination, hypokinesis and tremors. There are statistically significant differences in the rotarod test, alterations in

gait, and hindlimb clasping. Mice show neuritic pathology characteristic of human HD. Unlike the Bates model, there is limited neuronal loss.

Detloff and colleagues created a mouse knock-in model with an extension of the endogenous mouse CAG repeat to approximately 150 CAGs. This model, the CHL2 line, shows more aggressive phenotypes than prior mouse knock-in models containing few repeats (Lin 2001). Measurable neurological deficits include clasping, gait abnormalities, nuclear inclusions and astrogliosis.

The present studies utilize the well-characterized Borchelt mouse model (N171-82Q, line 81), and the Detloff knock-in model, the CHL2 line. The initial targets for htt silencing were focused on sequences present in the N171-82Q transgene (exons 1-3). The use of this model was advantageous in the preliminary shRNA development because the RNAi search could focus on only the amino-terminal encoding sequences rather than the full length 14 kb mRNA. Figure 21 depicts the one-step cloning approach used to screen hairpins (Harper 2004). No effective shRNAs were found in exon 1, but several designed against exon 2, denoted shHDEx2.1 (5'-AAGAAAGAACTTTCAGCTACC-3', SEQ ID NO:91), shHDEx2.2 19 nt (5'- AGAACTTTCAGCTACCAAG - 3' (SEQ ID NO:92)), or shHDEx2.2 21 nt 5' -AAAGAACTTTCAGCTACCAAG - 3' (SEQ ID NO:93)) and exon 3 (shHDEx3.1 19 nt 5'-TGCCTCAACAAAGTTATCA-3' (SEQ ID NO:94) or shHDEx3.1 21 nt 5'-AATGCCTCAACAAAGTTATCA-3' (SEQ ID NO:95)) sequences were effective. In co-transfection experiments with shRNA expressing plasmids and the N171-82Q transcript target, shHDEx2.1 reduced N171-Q82 transcript levels by 80%, and protein expression by 60%.

In transient transfection assays shHDex2.1 did not silence a construct spanning exons 1-3 of mouse htt containing a 79 CAG repeat expansion, the mouse equivalent of N171-82Q. Next shHDEx2 into NIH 3T3 cells were transfected to confirm that endogenous mouse htt, which is expressed in NIH 3T3 cells, would not be reduced. Surprisingly, shHDEx2.1 and shHDEx3.1 silenced full-length mouse htt. In contrast, shHDEx2.2 silenced only the human N171-82Q transgene.

Yamamoto and colleagues and others have demonstrated that preformed inclusions can resolve (Yamamoto 2000). To test if RNAi could also reduce preformed aggregates, the inventors used a neuronal cell line, which, upon induction of Q80-eGFP expression, showed robust inclusion formation (Xia 2002). Cells laden with aggregates were mock-transduced, or transduced with recombinant virus expressing control shRNA, or shRNAs directed against GFP. The inventors found dramatic reduction in aggregates as assessed by fluorescence. Quantification showed dose dependent effects (Figure 22) that were corroborated by western blot (Xia 2002).

As indicated in Example 1 above, viral vectors expressing siRNAs can mediate gene silencing in the CNS (Xia 2002). Also, these studies were extended to the mouse model of spinocerebellar ataxia type 1 (SCA1). The data are important as they demonstrate that shRNA is efficacious in the CNS of a mouse model of human neurodegenerative disease. The data also support that shRNA expression in brain is not detrimental to neuronal survival.

shRNAs can target the Exon 58 polymorphism. As described in Example 4 above, a polymorphism in htt exon 58 is in linkage disequilibrium with HD (Ambrose 1994). Thirty eight percent of the HD population possesses a 3-GAG repeat in exon 58, in contrast to the 4-GAG repeat found in 92% of non-HD patients. The polymorphism likely has no affect on htt, but it provides a target for directing gene silencing to the disease allele. As indicated in Example 4 above, in experiments to test if allele-specific silencing for HD was possible, plasmids were generated that expressed shRNAs that were specific for the exon 58 polymorphism. The exon 58 3-GAG-targeting shRNAs were functional.

Developing vectors for control of RNAi in vivo. As demonstrated above, shRNA expressed from viral vectors is effective at directing gene silencing in brain. Also, viral vectors expressing shSCA1 inhibited neurodegeneration in the SCA1 mouse model. ShRNA expression was constitutive in both instances. However, constitutive expression may not be necessary, and could exacerbate any noted nonspecific effects. The present inventors have developed and tested several doxycycline-regulated constructs. The construct depicted in Figure 23

showed strong suppression of target gene (GFP) expression after addition of doxycycline and RNAi induction.

RNAi can protect, and/or reverse, the neuropathology in mouse models of human Huntington's disease

Two distinct but complimentary mouse models are used, the N171-82Q transgenic and CHL2 knock-in mice. The former express a truncated NH2-terminal fragment of human htt comprising exons 1-3 with an 82Q-repeat expansion. The knock-in expresses a mutant mouse allele with a repeat size of ~150. Neither shows significant striatal or cortical cell loss. Both therefore are suitable models for the early stages of HD. They also possess similarities in mid- and end-stage neuropathological phenotypes including inclusions, gliosis, and motor and behavioral deficits that will permit comparison and validation. On the other hand, the differences inherent in the two models provide unique opportunities for addressing distinct questions regarding RNAi therapy. For example, N171-82Q transgenic mice have relatively early disease onset. Thus efficacy can be assessed within a few months, in contrast to 9 months or more in the CHL2 line. Because the data showed that shHDEx2.2 targets the human transgene and not mouse HD, evaluate disease-allele specific silencing in N171-82Q mice is evaluated. In contrast, the CHL2 knock-in is important for testing how reducing expression of both the mutant and wildtype alleles impacts on the HD phenotype. Finally, both models should be investigated because any therapy for HD should be validated in two relevant disease models.

siRNA against human htt protects against inclusion formation in N171-82Q mice

The data show that it is possible to silence the human N171-82Q transgene in vitro, and work in reporter mice and SCA1 mouse models demonstrated efficacy of RNAi in vivo in brain. shHDEx2.2 constructs, expressed from two vector systems with well-established efficacy profiles in CNS, are now tested for their capacity to reduce mutant transgenic allele expression in vivo. Further, the impact of shHDEx2.2 on inclusion formation is assessed. Inclusions may not be pathogenic themselves, but they are an

important hallmark of HD and their presence and abundance correlates with severity of disease in many studies.

Recombinant feline immunodeficiency virus (FIV) and adeno-associated virus (AAV) expressing shHDs are injected into N171-82Q. The levels of shHDs expressed from FIV and AAV are evaluated, as is the ability to reduce htt mRNA and protein levels in brain, and subsequently affect inclusion formation.

Mice. N171-82Q mice developed by Borchelt and colleagues are used for these experiments (Shilling 1999, Shilling 2001). The colony was set up from breeders purchased from Jackson Laboratories (N171-82Q, line 81) and are maintained as described (Shilling 1999, Shilling 2001). F1 pups are genotyped by PCR off tail DNA, obtained when tagging weaned litters.

IC2 and EM48 have been used previously to evaluate N171-82Q transgene expression levels in brain by immuno-histochemistry (IHC) and western blot (Zhou 2003, Trottier 1995). EM48 is an antibody raised against a GST-NH2 terminal fragment of htt that detects both ubiquitinated and non-ubiquitinated htt-aggregates (Li 2000), and the IC2 antibody recognizes long polyglutamine tracts (Trottier 1995). By 4 weeks N171-82Q mice show diffuse EM48-positive staining in striata, hippocampus, cerebellar granule cells, and cortical layers IV and V (Shilling 1999, Shilling 2001). The present experiments focus on the striatum and cortex because they are the major sites of pathology in human HD. TUNEL positivity and GFAP immunoreactivity are also significant in striatal sections harvested from 3 month old N171-82Q mice (Yu 2003). At 4 months, punctate nuclear and cytoplasmic immunoreactivity is also seen (Yu 2003).

Viruses. It is difficult to directly compare the two viruses under study at equivalent doses; FIV is enveloped and can be concentrated and purified, at best, to titers of 5×10^8 infectious units/ml (iu/ml). FIV pseudotyped with the vesicular stomatitus glycoprotein (VSVg) are used because of its tropism for neurons in the striatum (Brooks 2002). In contrast, AAV is encapsidated and can be concentrated and purified to titers ranging from 1×10^9 to 1×10^{11} iu/ml, with 1×10^{10} titers on average. AAV serotype 5 is used because it is tropic for

neurons in striatum and cortex, our target brain regions. Other serotypes of AAV, such as AAV-1 may also be used to neurons in striatum and cortex. Also, it diffuses widely from the injection site (Alisky 2000, Davidson 2000). Ten-fold dilutions of FIV and AAV generally results in a greater than 10-fold drop in transduction efficiency, making comparisons at equal titers, and dose escalation studies, unreasonable. Thus, both viruses are tested at the highest titers routinely available to get a fair assessment of their capacities for efficacy in N171-82Q mice. All viruses express the humanized Renilla reniformis green fluorescent protein (hrGFP) reporter transgene in addition to the shRNA sequence (Figure 24). This provides the unique opportunity to look at individual, transduced cells, and to compare pathological improvements in transduced vs. untransduced cells.

Injections. Mice are placed into a David Kopf frame for injections. Mice are injected into the striatum (5 microliters; 100 nl/min) and the cortex (3 microliters; 75 nl/min) using a Hamilton syringe and programmable Harvard pump. The somatosensory cortex is targeted from a burr hole at –1.5 mm from Bregma, and 1.5 mm lateral. Depth is 0.5 mm. The striatum is targeted through a separate burr hole at +1.1 mm from Bregma, 1.5 mm lateral and 2 mm deep. Only the right side of the brain is injected, allowing the left hemisphere to be used as a control for transgene expression levels and presence or absence of inclusions.

Briefly, groups of 4 week-old mice heterozygous for the N171-82Q transgene and their age-matched wildtype littermates are injected with FIV (FIV groups are VSVg.FIV.shHDEx2.2, VSVg.FIVshlacZ, VSVg.hrGFP, saline) or AAV (AAV groups are AAV5.shHDEx2.2, AAV5shlacZ, AAV5hrGFP, saline) (n=18/group; staggered injections because of the size of the experiment). Names of shHDEx2.2 and shlacZ expressing viruses have been shortened from shlacZ.hrGFP, for example, to make it easier to read, but all vectors express hrGFP as reporter. Nine mice/group are sacrificed at 12 weeks of age to assess the extent of transduction (eGFP fluorescence; viral copy number/brain region), shRNA expression (northern for shRNAs, and inhibition of expression of the transgenic allele (QPCR and western blot). The remaining groups are sacrificed

at 5 months of age. This experimental set up is repeated (to n=6/group) to confirm results and test inter-experiment variability.

All mice in all groups are weighed bi-weekly (every other week) after initial weekly measurements. N171-82Q mice show normal weight gain up to approximately 6 weeks, after which there are significant differences with their wildtype littermates.

PCR Analyses. Brains are harvested from mice sacrificed at 12 weeks of age, and grossly evaluated for GFP expression to confirm transduction. The cortex and striatum from each hemisphere is dissected separately, snap frozen in liquid N2, pulverized with a mortar and pestle, and resuspended in Trizol (Gibco BRL). Separate aliquots are used for Q-RTPCR for N171-82Q transgenes and DNA PCR for viral genomes. A coefficient of correlation is determined for transgene silencing relative to viral genomes for both vector systems, for the regions analyzed and compared to contralateral striata and mice injected with control vectors or saline.

The RNA harvested is used to evaluate activation of interferon-responsive genes. Bridges et al (Bridges 2003) and Sledz and colleagues (Sledz 2003) found activation of 2'5' oligo(A) polymerase (OAS) in cell culture with siRNAs and shRNAs, the latter expressed from lentivirus vectors. Gene expression changes are assessed using QPCR for OAS, Stat1, interferon-inducible transmembrane proteins 1 and 2 and protein kinase R (PKR). PKR activation is an initial trigger of the signaling cascade of the interferon response.

Protein analyses. A second set of 3 brains/group are harvested for protein analysis. Regions of brains are micro dissected as described above, and after pulverization are resuspended in extraction buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% SDS, 1 mM BetaME, 1X complete protease inhibitor cocktail) for analysis by western blot. HrGFP expression are evaluated and correlated to diminished levels of soluble N171-82Q using anti-GFP and antibodies to the NH2-terminal region of htt (EM48) or the polyglutamine tract (IC2).

Histology. Histology is done on the remaining animals. Mice are perfused with 2% paraformaldehyde in PBS, brains blocked to remove the cerebellum, post-fixed ON, and then cryoprotected in 30% sucrose. Full coronal sections (40 μm) of the entire cerebrum are obtained using a Microtome (American Products Co #860 equipped with a Super Histo Freeze freezing stage). Briefly, every section is collected, and sections 1-6 are placed into 6 successive wells of a 24-well plate. Every 400 microns, two sections each of 10 microns are collected for Nissl and H&E staining. The process is repeated.

EM-48 immuno-staining reveals diffuse nuclear accumulations in N171-82Q mice as early as 4 weeks of age. In 6 mo. old mice inclusions are extensive (Shilling 2001). The increase in cytoplasmic and nuclear EM48 immuno-reactivity, and in EM48 immuno-reactive inclusions over time allow quantitative comparisons between transduced and untransduced cells. Again, control values are obtained from mice injected with shlacZ-expressing vectors, saline injected mice, and wt mice. The contralateral region is used as another control, with care taken to keep in mind the possibility of retrograde and anterograde transport of virus from the injection site.

Quantitation of nuclear inclusions is done using BioQuant™ software in conjunction with a Leitz DM RBE upright microscope equipped with a motorized stage (Applied Scientific Instruments). Briefly, floating sections are stained with anti-NeuN (AMCA secondary) and EM48 antibodies (rhodamine secondary) followed by mounting onto slides. The regions to be analyzed are outlined, and threshold levels for EM48 immunoreactivity set using sections from control injected mice. A minimum of 50 hrGFP-positive and hrGFP negative neurons cells are evaluated per slide (5 slides/mouse), and inclusion intensity measured (arbitrary units). This is done for both striata and cortices. To quantitate cytoplasmic inclusions, the striatum is outlined and total EM48 aggregate density measured. Threshold values are again done using control hemispheres and control injected mice.

Additional wells of sections are stained with anti-GFAP, anti-neurofilament, and the lectin GSA to assay for viral or viral + hairpin induced

gliosis, neuritic changes, and microglial activation, respectively. GFAP-stained brain sections from N171-82Q mice show gliosis by 4 months (Yu 1998), although earlier time points have not been reported.

Stereology. In a separate experiment on N171-82Q mice and wt mice, unbiased stereology using BioQuant™ software is done to assess transduction efficiency. Stereology allows for an unbiased assessment of efficiency of transduction (number of cells transduced/input). AAV5 (AAV5hrGFP, AAV5shHD.hrGFP) and FIV (VSVg.FIVhrGFP, VSVg.FIVshHD.hrGFP) transduction efficiency is compared in the striatum and somatosensory cortex in HD and wildtype mice, with n=5 each. Mice are harvested at 12 and 20 weeks. The cerebrum is sectioned in its entirety and stored at –20°C until analysis. Briefly, six weeks after gene transfer with VSVg.FIVhrGFP (n=3) or AAV5hrGFP (n=3), every section of an HD mouse cerebrum is mounted and an initial assessment of the required numbers of sections and grid and dissector size done using the coefficient of error (as determined by Martheron's quadratic approximation formula) as a guide.

The 171-82Q HD mouse model has important neuropathological and behavioral characteristics relevant to HD. Onset of disease occurs earlier than HD knock-in or YAC transgenic models, allowing an initial, important assessment of the protective effects of RNAi on the development of neuropathology and dysfunctional behavior, without incurring extensive long term housing costs. Admittedly, disease onset is slower and less aggressive than the R6/2 mice created by Bates and colleagues (Mangiarini 1996), but the R6/2 line is difficult to maintain and disease is so severe that it may be less applicable and less predicative of efficacy in clinical trials.

N171-82Q mice (n=6/group) and age-matched littermates (n=6/group) are be weighed twice a month from 4 wks on, and baseline rotarod tests performed at 5 and 7 weeks of age. Numbers of mice per group are as described in Schilling et al (Shilling 1999) in which statistically significant differences between N171-82Q and wildtype littermates were described. At 7 weeks of age (after testing is complete), AAV (AAVshHDEx2.2, AAVshlacZ, AAVhrGFP,

saline) or FIV (FIVshEx2.2, FIVshlacZ, FIVhrGFP, saline) is injected bilaterally into the striatum and cortex. Rotarod tests are repeated at 3-week intervals starting at age 9 weeks, until sacrifice at 6 months. The clasping behavior is assessed monthly starting at 3 months.

Behavioral testing. N171-82Q mice are given four behavioral tests, all of which are standard assays for progressive disease in HD mouse models. The tests allow comparisons of behavioral changes resulting from RNAi to those incurred in HD mouse models given other experimental therapies. For example, HD mice given cystamine or creatine therapy showed delayed impairments in rotarod performance, and in some cases delayed weight loss (Ferrante 2000, Dedeoglu 2002, Dedeogu 2003) In addition to the rotarod, which is used to assay for motor performance and general neurological dysfunction, the activity monitor allows assessment of the documented progressive hypoactivity in N171-82Q mice. The beam analysis is a second test of motor performance that has also been used in HD mice models (Carter 1999). Clasping, a phenotype of generalized neurological dysfunction, is straightforward and takes little time. Clasping phenotypes were corrected in R. Hen's transgenic mice possessing an inducible mutant htt.

Accelerated rotarod. N171-82Q and age-matched littermates are habituated to the rotarod at week 4, and 4 trials per day for 4 days done on week 5 and 7, and every 3 weeks hence using previously described assays (Shilling 1999, Clark 1997) in use in the lab. Briefly, 10 min trials are run on an Economex rotarod (Columbus Instruments) set to accelerate from 4 to 40 rpm over the course of the assay. Latency to fall is recorded and averages/group determined and plotted. Based on prior work (Shilling 1999) 6 mice will give sufficient power to assess significance.

Clasping behavior. Normal mice splay their limbs when suspended, but mice with neurological deficits can exhibit the opposite, with fore and hind limbs crunched into the abdomen (clasping). All mice are suspended and scored for clasping monthly. The clasp must be maintained for at least 30 sec. to be scored positive.

Activity monitor. Most HD models demonstrate hypokinetic behavior, particularly later in the disease process. This can be measured in several ways. One of the simplest methods is to monitor home cage activity with an infrared sensor (AB-system 4.0, Neurosci Co., LTD). Measurements are taken over 3 days with one day prior habituation to the testing cage (standard 12-hour light/dark cycle). Activity monitoring is done at 12, 17, and 20 and 23 weeks of age.

Beam walking. N171Q-82Q and age matched littermates are assayed for motor performance and coordination using a series of successively more difficult beams en route to an enclosed safety platform. The assay is as described by Carter et al (Carter 1999). Briefly, 1 meter-length beams of 28, 17 or 11 mm diameter are placed 50 cm above the bench surface. A support stand and the enclosed goal box flank the ends. Mice are trained on the 11 mm beam at 6 weeks of age over 4 days, with 3 trials per day. If mice can traverse the beam in < 20 sec. trials are initiated. A trial is then run on each beam, largest to smallest, with a 60 sec cutoff/beam and one minute rest between beams. A second trial is run and the mean scores of the two trials evaluated.

RNAi cannot replace neurons; it only has the potential to protect non-diseased neurons, or inhibit further progression of disease at a point prior to cell death. N171-82Q mice do not show noticeable cellular loss, and is therefore an excellent model of early HD in humans. The general methodology is the similar to that described above, except that the viruses are injected at 4 months, when N171-82Q mice have measurable behavioral dysfunction and inclusions. Animals are sacrificed at end stage disease or at 8 months, whichever comes first. Histology, RNA and protein in harvested brains are analyzed as described above.

It is important to confirm the biological effects of virally expressed shHDs in a second mouse model, as it is with any therapy. The Detloff knock-in mouse (the CHL2 line, also notated as HdhCAGQ150) is used as a second model of early HD disease phenotypes. These mice have a CAG expansion of approximately 150 units, causing brain pathologies similar to HD including

gliosis and neural inclusions in the cortex and striatum. They also show progressive motor dysfunction and other behavioral manifestations including rotarod deficits, clasping, gait abnormalities and hypoactivity.

Heterozygous CHL2 mice express the mutant and wildtype allele at roughly equivalent levels, and shRNAs directed against mouse HD silence both transcripts. shmHDEx2.1 causes reductions in gene expression, but not complete silencing. Disease severity in mouse models is dependent on mutant htt levels and CAG repeat length.

The inventors created shmHDEx2 (shRNA for murine HD) directed against a region in mouse exon 2 that reduces expression of the full-length mouse Hdh transcript in vitro. Transduction of neurons with shmHDEx2-expressing viruses, and its impacts on neuropathological progression, behavioral dysfunction and the appearance of EM48 immuno-reactive inclusions in CHL2 mice is tested. shmHD-or shlacZ-expressing vectors in CHL2 and wildtype brain is tested. In this experiment, virus is injected into the striatum of wt or CHL2 mice (10/group) using the coordinates described above, at 3 months of age. Two months later mice are sacrificed and brains removed and processed for RNA (n=5/group) and protein (n=5).

A second study tests the vectors in the Detloff model. Briefly, 15 mice per group are injected into the striatum and cortex at 3 months of age with AAV (AAVshmHD, AAVshlacZ, AAVhrGFP, saline) or FIV (VSVg.FIV.shmHD, VSVg.FIVshlacZ, VSVg.FIVhrGFP, saline) expressing the transgenes indicated. To assess the impact of RNAi, activity performed. The mice are sacrificed at 16-18 months of age and five brains/group are processed for histology and sections banked in 24-well tissue culture plates. The remaining brains are processed for RNA (n=6) and protein analysis (n=5). Northern blots or western blots are required to analyze wildtype and mutant htt expression because the only distinguishing characteristic is size.

Development of effective allele-specific siRNAs

Mutant htt leads to a toxic gain of function, and inhibiting expression of the mutant allele has a profound impact on disease (Yamamoto 2000). Also, selectively targeting the disease allele would be desirable if non-disease allele silencing is deleterious. At the present time, there is one documented disease linked polymorphism in exon 58 (Lin 2001). Most non-HD individuals have 4 GAGs in Hdh exon 58 while 38% of HD patients have 3 GAGs. As described above, RNAi can be accomplished against the 3-GAG repeat.

Prior work by the inventors showed the importance of using full-length targets for testing putative shRNAs. In some cases, shRNAs would work against truncated, but not full-length targets, or vice-versa. Thus, it is imperative that testable, full-length constructs are made to confirm allele-specific silencing. The V5 and FLAG tags provide epitopes to evaluate silencing at the mRNA and protein levels. This is important as putative shRNAs may behave as miRNAs, leading to inhibition of expression but not message degradation.

Designing the siRNAs. Methods are known for designing siRNAs (Miller 2003, Gonzalez-Alegre 2003, Xia 2002, Kao 2003). Information is also know about the importance of maintaining flexibility at the 5' end of the antisense strand for loading of the appropriate antisense sequence into the RISC complex (Khvorova 2003 Schwarz 2003). DNA sequences are generated by PCR. This method allows the rapid generation of many candidate shRNAs, and it is significantly cheaper than buying shRNAs. Also, the inserts can be cloned readily into our vector shuttle plasmids for generation of virus. The reverse primer is a long oligonucleotide encoding the antisense sequence, the loop, the sense sequence, and a portion of the human U6 promoter. The forward primer is specific to the template in the PCR reaction. PCR products are cloned directly into pTOPO blunt from InVitrogen, plasmids transformed into DH5a, and bacteria plated onto Kanr plates (the PCR template is Ampr). Kanr clones are picked and sequenced. Sequencing is done with an extended 'hot start' to allow effective read-through of the hairpin. Correct clones are transfected into cells along with plasmids expressing the target or control sequence

(HttEx58.GAG3V5 and HttEx58.GAG4FLAG, respectively) and silencing evaluated by western blot. Reductions in target mRNA levels are assayed by Q-RTPCR. The control for western loading is neomycin phosphotransferase or hrGFP, which are expressed in the target-containing plasmids and provide excellent internal controls for transfection efficiency. The control for Q-RTPCR is HPRT.

Cell lines expressing targets with the identified polymorphism or control wildtype sequences are created. Target gene expression are under control of an inducible promoter. PC6-3, Tet repressor (TetR+) cells, a PC-12 derivative with a uniform neuronal phenotype (Xia 2002) are used. PC6-3 cells are transfected with plasmids expressing HDEx58.GAG3V5 (contains neo marker) and HDEx58GAG4FLG (contains puro marker), and G418+/puromycin+ positive clones selected and characterized for transcript levels and htt-V5 or htt-Flag protein levels.

FIV vectors expressing the allele specific shRNAs are generated and used to test silencing in the inducible cell lines. FIV vectors infect most epithelial and neuronal cell lines with high efficiency and are therefore useful for this purpose. They also efficiently infect PC6-3 cells. AAV vectors are currently less effective in in vitro screening because of poor transduction efficiency in many cultured cell lines.

Cells are transduced with 1 to 50 infectious units/cell in 24-well dishes, 3 days after induction of mutant gene expression. Cells are harvested 72 h after infection and the effects on HDEx58.GAG3V5 or HDEx58GAG4FLG expression monitored.

Example 7
Micro RNAi-Therapy for Polyglutamine Disease

Post-transcriptional gene silencing occurs when double stranded RNA (dsRNA) is introduced or naturally expressed in cells. RNA interference (RNAi) has been described in plants (quelling), nematodes, and *Drosophila*. This

process serves at least two roles, one as an innate defense mechanism, and another developmental (Waterhouse 2001 Fire 1999, Lau 2001, Lagos-Quintana 2001, Lee 2001). RNAi may regulate developmental expression of genes via the processing of small, temporally expressed RNAs, also called microRNAs (Knight 2001, Grishok 2001). Harnessing a cell's ability to respond specifically to small dsRNAs for target mRNA degradation has been a major advance, allowing rapid evaluation of gene function (Gonczy 2000, Fire 1998, Kennerdell 1998, Hannon 2002, Shi 2003, Sui 2002).

Most eukaryotes encode a substantial number of small noncoding RNAs termed micro RNAs (miRNAs) (Zeng 2003, Tijsterman 2004, Lee 2004, Pham 2004). mir-30 is a 22-nucleotide human miRNA that can be naturally processed from a longer transcript bearing the proposed miR-30 stem-loop precursor. mir-30 can translationally inhibit an mRNA-bearing artificial target sites. The mir-30 precursor stem can be substituted with a heterologous stem, which can be processed to yield novel miRNAs and can block the expression of endogenous mRNAs.

Huntington's disease (HD) and Spinocerebellar ataxia type I (SCA1) are two of a class of dominant, neurodegenerative diseases caused by a polyglutamine (polyQ) expansion. The mutation confers a toxic gain of function to the protein, with polyQ length predictive of age of onset and disease severity. There is no curative or preventative therapy for HD or SCA1, supporting the investigation of novel strategies. As described above, the inventors showed that gene silencing by RNA interference (RNAi) can be achieved in vitro and in vivo by expressing short hairpin RNAs (shRNAs) specific for mRNAs encoding ataxin-1 or huntingtin. Currently, strong, constitutive po1III promoters (U6 and H1) are used to express shRNAs, which are subsequently processed into functional small interfering RNAs (siRNAs). However, strong, constitutive expression of shRNAs may be inappropriate for diseases that take several decades to manifest. Moreover, high-level expression may be unnecessary for sustained benefit, and in some systems may induce a non-specific interferon response leading to global shut-down of gene expression. The inventors

therefore generated polII-expressed microRNAs (miRNAs) as siRNA shuttles as an alternative strategy. Due to their endogenous nature, miRNA backbones may prevent the induction of the interferon response.

Using human mir-30 as a template, miRNA shuttles were designed that upon processing by dicer released siRNAs specific for ataxin-l. Briefly, the constructs were made by cloning a promoter (such as an inducible promoter) and an miRNA shuttle containing an embedded siRNA specific for a target sequence (such as ataxin-1) into a viral vector. By cloning the construct into a viral vector, the construct can be effectively introduced in vivo using the methods described in the Examples above. Constructs containing po1II-expressed miRNA shuttles with embedded ataxin-1-specific siRNAs were co-transfected into cells with GFP-tagged ataxin-l, and gene silencing was assessed by fluorescence microscopy and western analysis. Dramatic arid dose-dependent gene silencing relative to non-specific miRNAs carrying control siRNAs was observed. This po1II-based expression system exploits the structure of known miRNAs and supports tissue-specific as well as inducible siRNA expression, and thus, serves as a unique and powerful alternative to dominant neurodegenerative disease therapy by RNAi.

Briefly, the constructs were made by cloning a promoter (such as an inducible promoter) and an miRNA shuttle containing an embedded siRNA specific for a target sequence (such as ataxin-1) into a viral vector. By cloning the construct into a viral vector, the construct can be effectively introduced in vivo using the methods described in the Examples above.

All publications, patents and patent applications are incorporated herein by reference. While in the foregoing specification this invention has been described in relation to certain embodiments thereof, and many details have been set forth for purposes of illustration, it will be apparent to those skilled in the art that the invention is susceptible to additional embodiments and that certain of the details described herein may be varied considerably without departing from the basic principles of the invention.

Citations

Adelman *et al.*, DNA, 2, 183 (1983).

Alisky *et al.*, Hum Gen Ther, 11,2315 (2000b).

Alisky *et al.*, NeuroReport, 11,2669 (2000a).

Altschul *et al.*, JMB, 215, 403 (1990).

Altschul *et al.*, Nucleic Acids Res. 25, 3389 (1997).

Ambrose *et al.*, Somat Cell Mol Genet.20, 27-38 (1994)

Anderson *et al.*, Gene Ther., 7(12), 1034-8 (2000).

Andreason and Evans, Biotechniques, 6, 650 (1988).

Augood *et al*,. Neurology, 59, 445-8 (2002).

Augood *et al.*, Ann. Neurol., 46, 761-769 (1999).

Bass, Nature, 411, 428 (2001).

Batzer *et al.*, Nucl. Acids Res.,19, 508 (1991).

Baulcombe, Plant Mol. Biol., 32, 79 (1996).

Bates *et al.*, Curr Opin Neurol 16:465-470, 2003.

Behr *et al.*, Proc. Natl. Acad. Sci. USA, 86, 6982 (1989).

Bernstein *et al.*, Nature, 409, 363 (2001).

Bledsoe *et al.*, NatBiot, 18, 964 (2000).

Brantl, Biochemica and Biophysica Acta, 1575, 15 (2002).

Brash *et al.*, Molec. Cell. Biol., 7, 2031 (1987).

Breakefield *et al.*, Neuron, 31, 9-12 (2001).

Bridge *et al.*, Nat Genet 34:263-264, 2003.

Brooks *et al.*, Proc. Natl. Acad. Sci. U. S. A., 99,6216 (2002).

Brummelkamp, T.R. *et al.*, *Science* **296**:550-553 (2002).

Burright, E.N. *et al.*, Cell, 82, 937-948 (1995)

Capecchi, Cell, 22, 479 (1980).

Caplan *et al.*, Proc. Natl. Acad. Sci. U. S. A., 98, 9742 (2001).

Caplen *et al.*, Hum. Mol. Genet., 11(2), 175-84 (2002).

Carter *et al.*, J Neurosci 19:3248, 1999.

Cemal *et al.*, Hum. Mol. Genet., 11(9), 1075-94 (2002).

Chai *et al.*, Hum. Mol. Genet., 8, 673-682 (1999b).

Chai *et al.*, J. Neurosci., 19, 10338 (1999).

Chan *et al.*, Hum Mol Genet., 9(19), 2811-20 (2000).

Chen, H.K. *et al.*, Cell, 113, 457-68 (2003)

Chiu and Rana, Mol. Cell., 10(3), 549-61 (2002).

Clark, H.B. *et al.*, J. Neurosci., 17(19), 7385-7395 (1997)Cogoni *et al.*, Antonie Van Leeuwenhoek, 65, 205 (1994).

Corpet *et al.*, Nucl. Acids Res., 16, 10881 (1988).

Crea *et al.*, Proc. Natl. Acad. Sci. U.S.A., 75, 5765 (1978).

Cullen, Nat. Immunol., 3, 597-9 (2002).

Cummings, C.J. *et al.*, Nat. Genet., 19(2), 148-154 (1998)

Davidson *et al.*, Proc. Natl. Acad. Sci. U. S. A., 97, 3428 (2000).

Davidson, B.L. *et al.*, The Lancet Neurol., 3, 145-149 (2004)

Davidson *et al.*, Nat Rev Neurosci 4:353-364, 2003.

Dayhoff *et al.*, Atlas of Protein Sequence and Structure (Natl. Biomed. Res. Found. 1978)

Dedeoglu *et al.*, J Neurochem 85:1359-1367, 2003.

Dedeoglu *et al.*, J Neurosci 22:8942-8950, 2002.

Doheny *et al.*, Neurology, 59, 1244-1246 (2002).

Donze and Picard, Nucleic Acids Res., 30(10) (2002).

During *et al.*, Gene Ther 5:820-827, 1998.

Elbashir *et al.*, EMBO J., 20(23), 6877-88 (2001c).

Elbashir *et al.*, Genes and Development, 15, 188 (2001).

Elbashir *et al.*, Nature, 411, 494 (2001).

Emamian, E.S. *et al.*, Neuron, 38, 375-87 (2003)

Fahn *et al.*, Adv. Neurol., 78, 1-10 (1998).

Felgner *et al.*, Proc. Natl. Acad. Sci., 84, 7413 (1987).

Fernandez-Funez, P. *et al.*, Nature, 408, 101-106 (2000)

Ferrante *et al.*, J Neurosci 20:4389-4397, 2000.

Fire *et al.*, Nature, 391(6669), 806-11 (1998).

Fire A. Trends Genet 15(9):358-363, 1999

Frisella *et al.*, Mol Ther 3(3):351-358, 2001.

Gaspar *et al.*, Am. J. Hum. Genet., 68(2), 523-8 (2001).

Gelfand, PCR Strategies, Academic Press (1995).

Gitlin *et al.*, Nature, 418(6896), 430-4 (2002).

Goeddel *et al.*, Nucleic Acids Res., 8, 4057 (1980).

Gonczy *et al.*, Nature 408:331-336, 2000.

Gonzalez-Alegre *et al.*, Nat Genet 3:219-223, 1993. Goodchild *et al.*, Mov. Disord., 17(5), 958, Abstract (2002).

Grishok *et al.*, Cell 106:23-34, 2001.

Hamilton and Baulcombe, Science, 286, 950 (1999).

Hammond *et al.*, Nature, 404, 293 (2000).

Harper *et al.*, Meth Mol Biol. In Press 2004

Hewett *et al.*, Hum. Mol. Gen., 9, 1403-1413 (2000).

Higgins *et al.*, CABIOS, 5, 151 (1989).

Higgins *et al.*, Gene, 73, 237 (1988).

Hilberg *et al.*, Proc. Natl. Acad. Sci. USA, 84, 5232 (1987).

Holland *et al.*, Proc. Natl. Acad. Sci. USA, 84, 8662 (1987).

Hornykiewicz *et al.*, N. Engl. J. Med,. 315, 347–353 (1986).

Huang *et al.*, CABIOS, 8, 155 (1992).

Hutton *et al.*, Nature, 393, 702-705 (1998).

Innis and Gelfand, PCR Methods Manual, Academic Press (1999).

Innis *et al.*, PCR Protocols, Academic Press (1995).

Jacque *et al.*, Nature, 418(6896), 435-8 (2002).

Johnston, Nature, 346, 776 (1990).

Kang *et al.*, J Virol 76:9378-9388, 2002.

Karlin and Altschul, Proc. Natl. Acad. Sci. USA, 87, 2264 (1990).

Karlin and Altschul, Proc. Natl. Acad. Sci. USA, 90, 5873 (1993)

Kennerdell and Carthew, Cell, 95, 1017 (1998).

Kao *et al.*, J Biol Chem 2003.

Kawasaki, H., *et al.*, Nucleic Acids Res, 31, 981-7 (2003)

Khvorova, A., *et al.*, Cell, 115, 505 (2003)

Kitabwalla and Ruprecht, N. Engl. J. Med., 347, 1364-1367 (2002).

Klein *et al.*, Ann. Neurol., 52, 675-679 (2002).

Klein *et al.*, Curr. Opin. Neurol., 4, 491-7 (2002).

Klement, I.A. *et al.*, Cell, 95, 41-53 (1998)

Knight *et al.*, Science 293:2269-2271, 2001.

Konakova *et al.*, Arch. Neurol., 58, 921-927 (2001).

Krichevsky and Kosik, Proc. Natl. Acad. Sci. U.S.A., 99(18),11926-9 (2002).

Kriegler, M. Gene Transfer and Expression, A Laboratory Manual, W.H. Freeman Co, New York, (1990).

Kunath *et al.*, Nat Biotechnol 21:559-561, 2003.

Kunkel *et al.*, Meth. Enzymol., 154, 367 (1987).

Kunkel, Proc. Natl. Acad. Sci. USA, 82, 488 (1985).

Kustedjo *et al.*, J. Biol. Chem., 275, 27933-27939 (2000).

Laccone *et al.*, Hum. Mutat., 13(6), 497-502 (1999).

Lagos-Quintana *et al.*, Science 294:853-858, 2001.

Lai *et al.*, Proc. Natl. Acad. Sci. USA, 86, 10006 (1989).

Larrick, J. W. and Burck, K. L., Gene Therapy. Application of Molecular Biology, Elsevier Science Publishing Co., Inc., New York, p. 71-104 (1991).

Lau *et al.*, Science 294:858-862, 2001.

Lawn *et al.*, Nucleic Acids Res., 9, 6103 (1981).

Lee, N.S., *et al.*, *Nat. Biotechnol.* **19**:500-505 (2002).

Lee *et al.*, Science 294:862-864, 2001.

Lee *et al.*, Cell, 117, 69-81 (2004)

Leger *et al.*, J. Cell. Sci., 107, 3403-12 (1994).

Leung *et al.*, Neurogenetics, 3, 133-43 (2001).

Li *et al.*, Nat Genet 25:385-389, 2000.

Lin *et al.*, Hum. Mol. Genet., 10(2), 137-44 (2001).

Loeffler *et al.*, J. Neurochem., 54, 1812 (1990).

Lotery *et al.*, Hum Gene Ther 13:689-696, 2002.

Mangiarini *et al.*, Cell 87(3):493-506, 1996.

Manche *et al.*, Mol. Cell Biol., 12, 5238 (1992).

Margolis and Ross, Trends Mol. Med., 7, 479 (2001).

Martinez *et al.*, Cell, 110(5), 563-74 (2002).

McCaffrey *et al.*, Nature, 418(6893), 38-9 (2002).

McManus and Sharp, Nat. Rev. Genet. 3(10), 737-47 (2002).

Meade *et al.*, J Comp Neurol 449:241-269, 2002.

Meinkoth and Wahl, Anal. Biochem., 138, 267 (1984).

Methods in Molecular Biology, 7, Gene Transfer and Expression Protocols, Ed. E. J. Murray, Humana Press (1991).

Miller, *et al.*, Mol. Cell. Biol., 10, 4239 (1990).

Miller, V.M. *et al.*, PNAS USA, 100, 7195-200 (2003)

Minks *et al.*, J. Biol. Chem., 254, 10180 (1979).

Miyagishi, M. & Taira, K. *Nat. Biotechnol.* **19**:497-500 (2002).

Moulder *et al.*, J. Neurosci., 19, 705 (1999).

Murray, E. J., ed. Methods in Molecular Biology, Vol. 7, Humana Press Inc., Clifton, N.J., (1991).

Myers and Miller, CABIOS, 4, 11 (1988).

Nasir *et al.*, Cell, 81, 811-823 (1995).

Needleman and Wunsch, JMB, 48, 443 (1970).

Nykänen *et al.*, Cell, 107, 309 (2001).

Ogura and Wilkinson, Genes Cells, 6, 575-97 (2001).

Ohtsuka *et al.*, JBC, 260, 2605 (1985).

Okabe *et al.*, FEBS Lett., 407, 313 (1997).

Ooboshi *et al.*, Arterioscler. Thromb. Vasc. Biol., 17, 1786 (1997).

Orr *et al.*, Nat. Genet. 4, 221-226 (1993)

Orr *et al.*, Cell 101:1-4, 2000.

Ozelius *et al.*, Genomics, 62, 377-84 (1999).

Ozelius *et al.*, Nature Genetics, 17, 40-48 (1997).

Passini *et al.*, J Virol 77:7034-7040, 2003.

Paul, C.P., *et al.*, *Nat. Biotechnol.* **19**:505-508 (2002).

Paulson *et al.*, Ann. Neurol., 41(4), 453-62 (1997).

Pearson and Lipman, Proc. Natl. Acad. Sci. USA, 85, 2444 (1988).

Pearson *et al.*, Meth. Mol. Biol., 24, 307 (1994).

Pham *et al.*, Cell, 117:83-94 (2004)

Pittman *et al.*, J. Neurosci., 13(9), 3669-80 (1993).

Plasterk *et al.*, Cell, 117, 1-4 (2004)

Poorkaj *et al.*, Ann. Neurol., 43, 815-825 (1998).

Quantin, B., *et al.*, Proc. Natl. Acad. Sci. USA, 89, 2581 (1992).

Reynolds, A. *et al.*, Nat. Biotechnol,. 22, 326-30 (2004)

Rosenfeld, M. A., *et al.*, Science, 252, 431 (1991).

Rossolini *et al.*, Mol. Cell. Probes, 8, 91 (1994).

Rubinson *et al.*, Nat Genet 33:401-406, 2003.

Sambrook and Russell, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press Cold Spring Harbor, NY (2001).

Scharfmann *et al.*, Proc. Natl. Acad. Sci. USA, 88, 4626 (1991).

Schilling *et al.*, Hum Mol Genet 8(3):397-407, 1999.

Schilling *et al.*, Neurobiol Dis 8:405-418, 2001.

Schwarz *et al.*, Mol. Cell., 10(3), 537-48 (2002).

Shipley *et al.*, J. Biol. Chem., 268, 12193 (1993).

Skinner, P.J. *et al.*, Nature, 389, 971-234 (1997)

Skorupa *et al.*, Exp Neurol 160:17-27, 1999.

Sledz *et al.*, Nat Cell Biol 5:834-839, 2003.

Smith *et al.*, Adv. Appl. Math., 2, 482 (1981).

Stein *et al.*, J. Virol., 73, 3424 (1999).

Stein *et al.*, Mol Ther 3(6):850-856, 2001.

Stein *et al.*, RNA, 9(2), 187-192 (2003).

Svoboda *et al.*, Development, 127, 4147 (2000).

Tanemura *et al.*, J. Neurosci., 22(1), 133-41 (2002).

Tang *et al.*, Genes Dev., 17(1), 49-63 (2003).

Ternin, H., "Retrovirus vectors for gene transfer", in Gene Transfer, Kucherlapati R, Ed., pp 149-187, Plenum, (1986).

Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology Hybridization with Nucleic Acid Probes, part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays" Elsevier, New York (1993).

Timmons and Fire, Nature, 395, 854 (1998).

Trottier *et al.*, Nature, 378(6555), 403-6 (1995).

Turner *et al.*, Mol. Biotech., 3, 225 (1995).

Tuschl, Nat. Biotechnol., 20, 446-8 (2002).

Urabe, M., *et al.*, Hum. Gene Ther., 13, 1935-1943 (2002)

Valerio *et al.*, Gene, 84, 419 (1989).

Viera *et al.*, Meth. Enzymol., 153, 3 (1987).

Walker and Gaastra, Techniques in Mol. Biol. (MacMillan Publishing Co. (1983).

Walker *et al.*, Neurology, 58, 120-4 (2002).

Waterhouse *et al.*, Proc. Natl. Acad. Sci. U. S. A., 95, 13959 (1998).

Waterhouse *et al.*, Nature 411:834-842, 2001.

Wianny and Zernicka-Goetz, Nat. Cell Biol., 2, 70 (2000).

Xia *et al.*, Nat. Biotechnol., 19, 640 (2001).

Xia *et al.*, Nat. Biotechnol., 20(10), 1006-10 (2002).

Xiao *et al.*,. Exp Neurol 144:113-124, 1997.

Yamamoto *et al.*, Cell, 101(1), 57-66 (2000).

Yang *et al.*, Mol. Cell Biol., 21, 7807 (2001).

Yu *et al.*, Proc. Natl. Acad. Sci., 99, 6047-6052 (2002).

Yu *et al.*, J Neurosci 23:2193-2202, 2003.

Zamore *et al.*, Cell, 101, 25 (2000).

Zeng *et al.*, RNA, 9:112-123 (2003)

Zhou *et al.*, J Cell Biol 163:109-118, 2003.

Zoghbi and Orr, Annu. Rev. Neurosci., 23, 217-47 (2000).

Zoghbi *et al.*, Semin Cell Biol., 6, 29-35 (1995)

Zu, T. *et al.*, In Preparation (2004)

WHAT IS CLAIMED IS:

1. An AAV-1 expressed siRNA comprising an isolated first strand of RNA of 15 to 30 nucleotides in length and an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first or second strand comprises a sequence that is complementary to a nucleotide sequence encoding a mutant Huntington's Disease protein, wherein at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex under physiological conditions, and wherein the siRNA silences the expression of the nucleotide sequence encoding the mutant Huntington's Disease protein in the cell.

2. The siRNA of claim 1, wherein the first and/or second strand further comprises a 3' overhang region, a 5' overhang region, or both 3' and 5' overhang regions.

3. The siRNA of claim 2, wherein the overhang region or regions is from 1 to 10 nucleotides in length.

4. The siRNA of claim 1, wherein the first strand and the second strand are operably linked by means of an RNA loop strand to form a hairpin structure comprising a duplex structure and a loop structure.

5. The siRNA of claim 4, wherein the loop structure contains from 4 to 10 nucleotides.

6. The siRNA of claim 4,wherein the loop structure corresponds to SEQ ID NO:61 or SEQ ID NO:64.

7. The siRNA of claim 1,wherein the first strand corresponds to SEQ ID NO:60 and the second strand corresponds to SEQ ID NO:62.

8. A mammalian cell comprising an expression cassette encoding an isolated first strand of RNA corresponding to SEQ ID NO:60, SEQ ID NO:63, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, or SEQ ID NO:88, and encoding an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first or second strand comprises a sequence that is complementary to a nucleotide sequence encoding a Huntington's Disease protein (htt), wherein at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex under physiological conditions, and wherein the siRNA silences the expression of the Huntington's Disease gene in the cell.

9. The mammalian cell of claim 8, wherein the expression cassette further comprises a promoter.

10. The mammalian cell of claim 9, wherein the promoter is a regulatable promoter.

11. The mammalian cell of claim 9, wherein the promoter is a constitutive promoter.

12. The mammalian cell of claim 9, wherein the promoter is a CMV, RSV, pol II or pol III promoter.

13. The mammalian cell of claim 8, wherein the expression cassette further comprises a marker gene.

14. The mammalian cell of claim 8, wherein the expression cassette is contained in a vector.

15. The mammalian cell of claim 14, wherein the vector is an adenoviral, lentiviral, adeno-associated viral (AAV), poliovirus, HSV, or murine Maloney-based viral vector.

16. The mammalian cell of claim 14, wherein the vector is an AAV vector.

17. The mammalian cell of claim 8,wherein the first strand corresponds to SEQ ID NO:60 and the second strand corresponds to SEQ ID NO:62.

18. A small interfering RNA (siRNA) comprising an first strand of RNA corresponding to SEQ ID NO:60, SEQ ID NO:63, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, or SEQ ID NO:88, and a second strand of RNA of 15 to 30 nucleotides in length, wherein the first or second strand comprises a sequence that is complementary to a nucleotide sequence encoding a Huntington's Disease protein (htt), wherein at least 12 nucleotides of the first and second strands are complementary to each other and form an siRNA duplex under physiological conditions, wherein the duplex is between 15 and 30 base pairs in length, and wherein the siRNA silences the expression of the Huntington's Disease gene in the cell.

19. The siRNA of claim 18, wherein the first and/or second strand further comprise a 3' overhang region, a 5' overhang region, or both 3' and 5' overhang regions.

20. The siRNA of claim 19, wherein the overhang region or regions is from 1 to 10 nucleotides in length.

21. The siRNA of claim 18, wherein the first strand and the second strand are operably linked by means of an RNA loop strand to form a hairpin structure comprising a duplex structure and a loop structure.

22. The siRNA of claim 21, wherein the loop structure contains from 4 to 10 nucleotides.

23. The siRNA of claim 21, wherein the loop structure corresponds to SEQ ID NO:61 or SEQ ID NO:64.

24. A method of performing Huntington's Disease gene silencing in a mammal comprising administering to the mammal an expression cassette encoding an isolated first strand of RNA corresponding to SEQ ID NO:60, SEQ ID NO:63, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, or SEQ ID NO:88, and encoding an isolated second strand of RNA of 15 to 30 nucleotides in length, wherein the first or second strand comprises a sequence that is complementary to a nucleotide sequence encoding a Huntington's Disease protein (htt), wherein at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex under physiological conditions, and wherein the expression of the siRNA from the expression cassette silences the expression of the Huntington's Disease gene in the mammal.

25. The method of claim 24, wherein the expression cassette further comprises a promoter.

26. The method of claim 25 wherein the promoter is a regulatable promoter.

27. The method of claim 25 wherein the promoter is a constitutive promoter.

28. The method of claim 25 wherein the promoter is a CMV, RSV, pol II or pol III promoter.

29. The method of claim 25, wherein the expression cassette further comprises a marker gene.

30. The method of claim 25, wherein the expression cassette is contained in a viral vector.

31. The method of claim 30, wherein the viral vector is an adenoviral, lentiviral, adeno-associated viral (AAV), poliovirus, HSV, or murine Maloney-based viral vector.

32. The method of claim 30, wherein the vector is an AAV vector.

33. The method of claim 24, wherein the first strand corresponds to SEQ ID NO:60 and the second strand corresponds to SEQ ID NO:62.

34. An isolated RNA comprising SEQ ID NO:59 that functions in RNA interference of a sequence encoding a mutant Huntington's Disease protein (htt).

35. An isolated RNA duplex comprising a first strand of RNA corresponding to SEQ ID NO:60 and a second strand of RNA corresponding to SEQ ID NO:62.

36. The RNA duplex of claim 35, wherein the first and/or second strand further comprises a 3' overhang region, a 5' overhang region, or both 3' and 5' overhang regions.

37. The RNA duplex of claim 36, wherein the overhang region or regions is from 1 to 10 nucleotides in length.

38. The RNA duplex of claim 36, wherein the first strand and the second strand are operably linked by means of an RNA loop strand to form a hairpin structure comprising a duplex structure and a loop structure.

39. The RNA duplex of claim 38, wherein the loop structure contains from 4 to 10 nucleotides.

40. The mammalian cell of claim 38,wherein the loop structure corresponds to SEQ ID NO:61 or SEQ ID NO:64.

41. A vector comprising two expression cassettes, a first expression cassette comprising a nucleic acid encoding a first strand of an RNA duplex corresponding to SEQ ID NO:60 and a second expression cassette comprising a nucleic acid encoding a second strand of the RNA duplex corresponding to SEQ ID NO:62.

42. The vector of claim 41, wherein the vector is a viral vector.

43. The vector of claim 42, wherein the viral vector is an adenoviral, lentiviral, adeno-associated viral (AAV), poliovirus, HSV, or murine Maloney-based viral vector.

44. The vector of claim 42, wherein the vector is an AAV vector.

45. A vector comprising an expression cassette, wherein the expression cassette encodes a nucleic acid SEQ ID NO:59.

46. The vector of claim 45, wherein the vector is a viral vector.

47. The vector of claim 46, wherein the viral vector is an adenoviral, lentiviral, adeno-associated viral (AAV), poliovirus, HSV, or murine Maloney-based viral vector.

48. The vector of claim 47, wherein the vector is an AAV vector.

49. A mammalian cell comprising an isolated first strand of RNA corresponding to SEQ ID NO:60, SEQ ID NO:63, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, or SEQ ID NO:88, and an isolated second strand

of RNA of 15 to 30 nucleotides in length, wherein the first or second strand comprises a sequence that is complementary to a nucleotide sequence encoding a Huntington's Disease protein (htt), wherein at least 12 nucleotides of the first and second strands are complementary to each other and form a small interfering RNA (siRNA) duplex under physiological conditions, and wherein the siRNA silences the expression of the Huntington's Disease gene in the cell.

50. The mammalian cell of claim 49, wherein the first and/or second strand further comprises a 3' overhang region, a 5' overhang region, or both 3' and 5' overhang regions.

51. The mammalian cell of claim 50, wherein the overhang region or regions is from 1 to 10 nucleotides in length.

52. The mammalian cell of claim 49, wherein the first strand and the second strand are operably linked by means of an RNA loop strand to form a hairpin structure comprising a duplex structure and a loop structure.

53. The mammalian cell of claim 52, wherein the loop structure contains from 4 to 10 nucleotides.

54. The mammalian cell of claim 52, wherein the loop structure corresponds to SEQ ID NO:61 or SEQ ID NO:64.

55. The mammalian cell of claim 49,wherein the first strand corresponds to SEQ ID NO:60 and the second strand corresponds to SEQ ID NO:62.

56. The siRNA of claim 1, wherein the first or second strand comprises a sequence that is complementary to both a mutant and wild-type Huntington's disease allele, wherein the siRNA silences the expression of the nucleotide sequence encoding the mutant Huntington's Disease protein and wild-type Huntington's Disease protein in the cell.

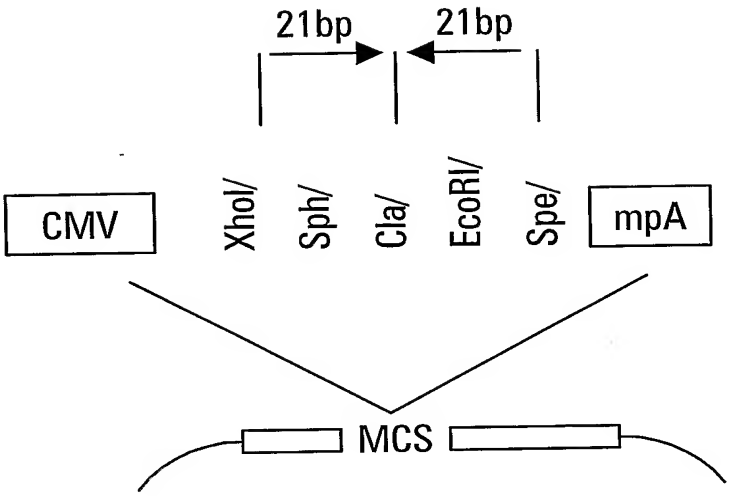


Fig. 1A

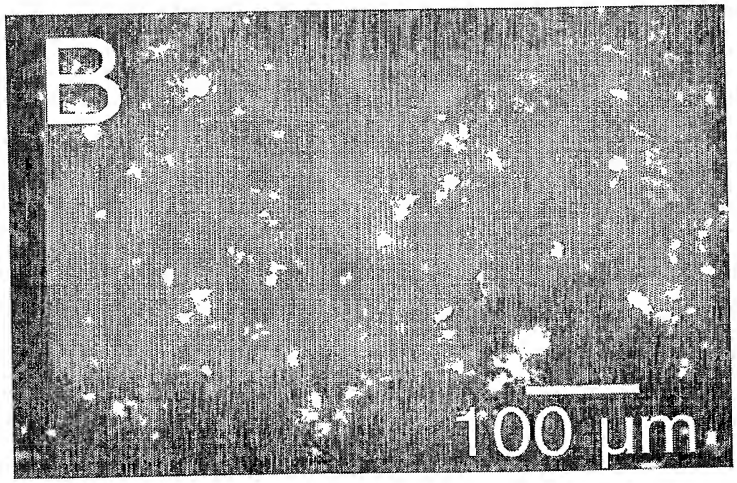


Fig. 1B

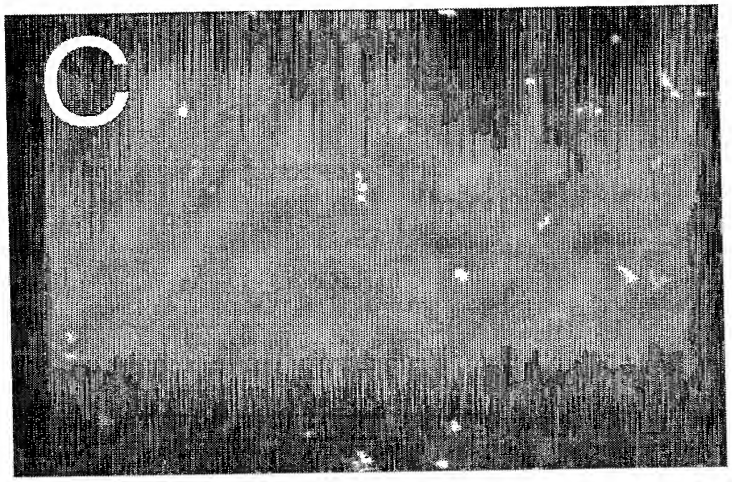


Fig. 1C

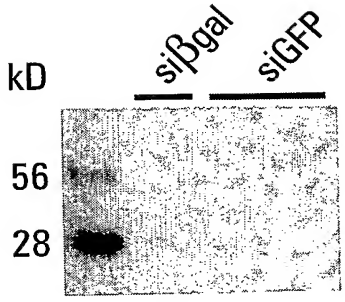


Fig. 1D

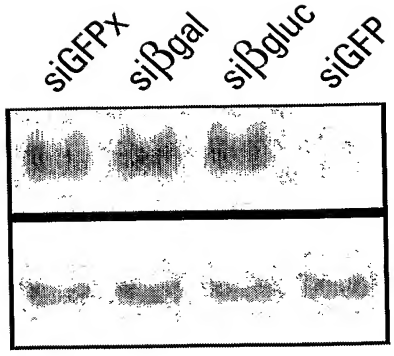


Fig. 1E

Fig. 1F

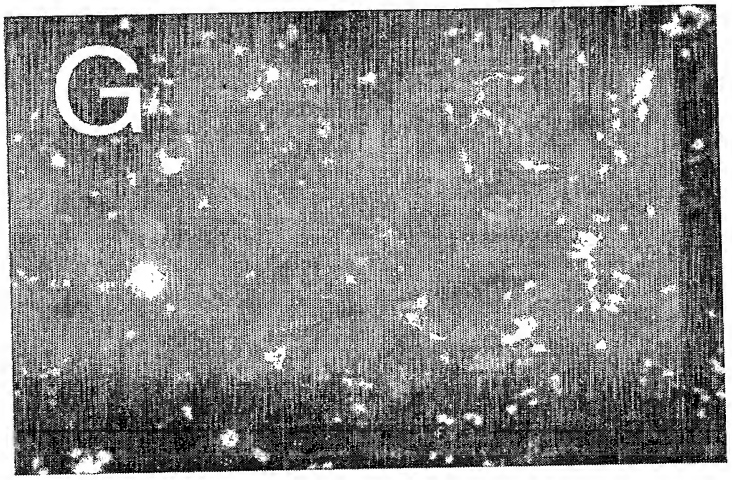


Fig. 1G

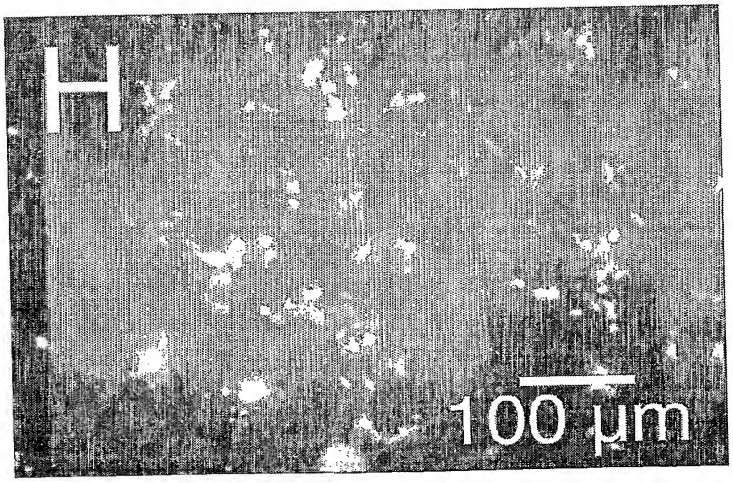


Fig. 1H

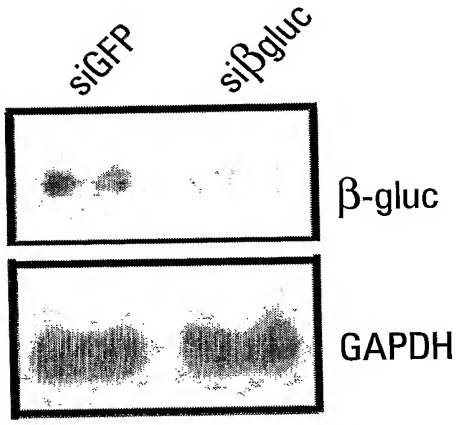

siGFP
siβgluc
β-gluc
GAPDH

Fig. 1I

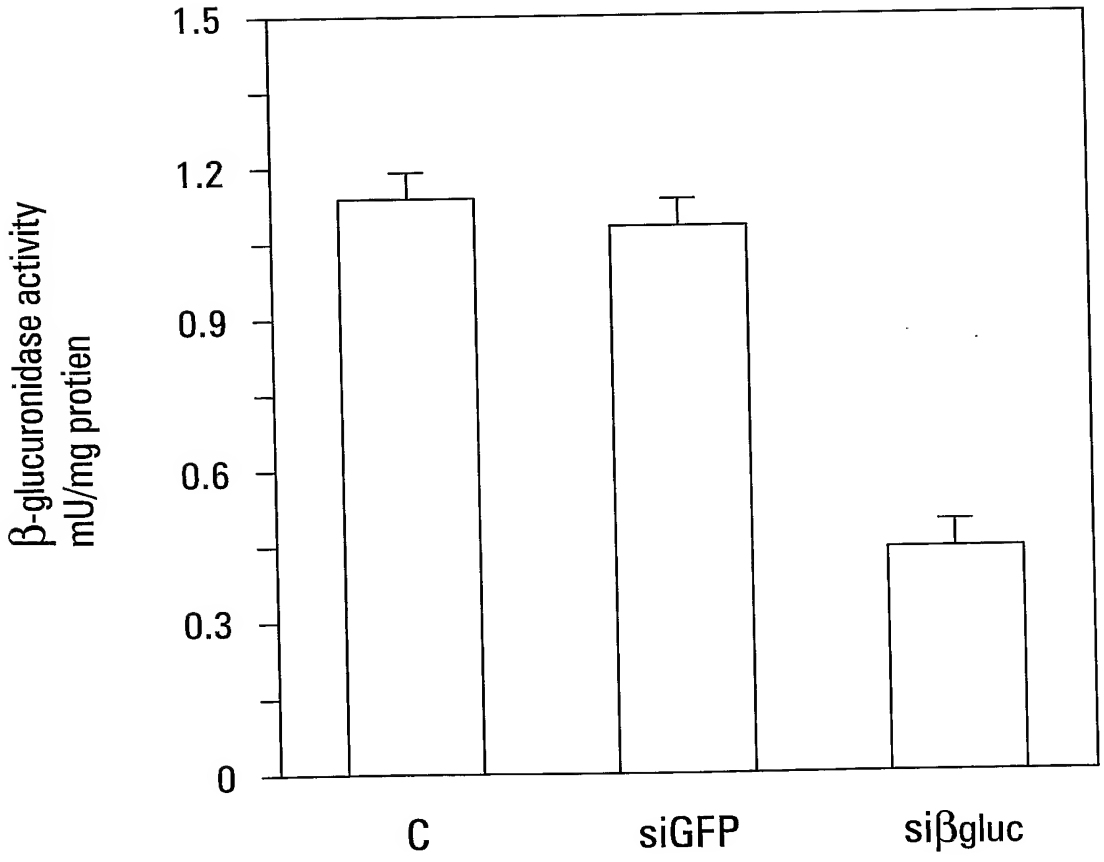
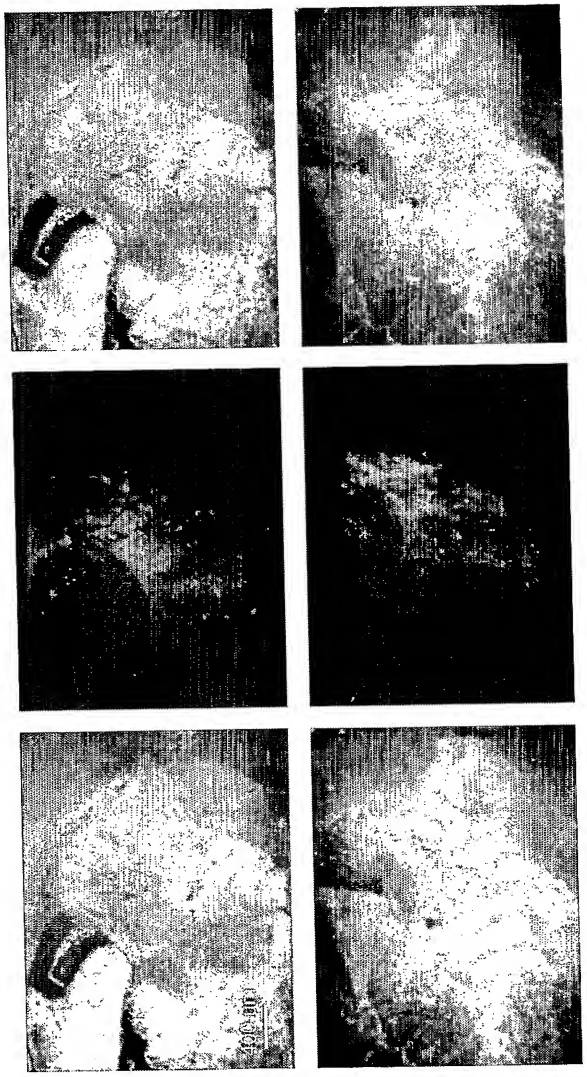

β-glucuronidase activity
mU/mg protien
1.5
1.2
0.9
0.6
0.3
0
C
siGFP
siβgluc

Fig. 1J



siGFP/dsRED siβgluc/dsRED

Fig. 2A

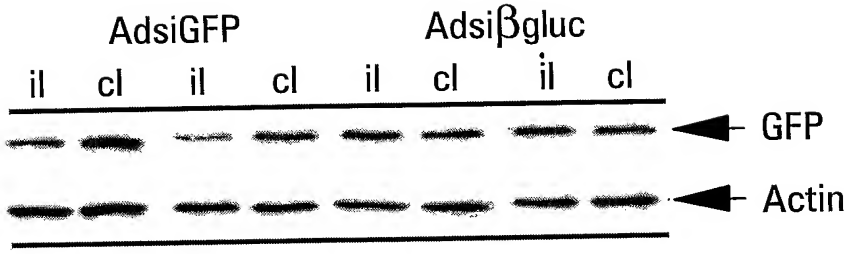

AdsiGFP
Adsiβgluc
il
cl
il
cl
il
cl
il
cl
GFP
Actin

Fig. 2B

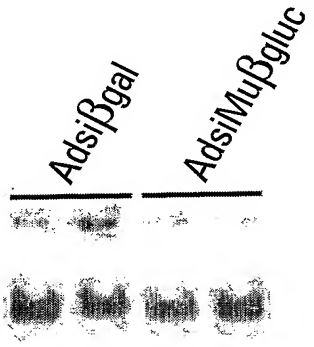

Adsiβgal
AdsiMuβgluc

Fig. 2C

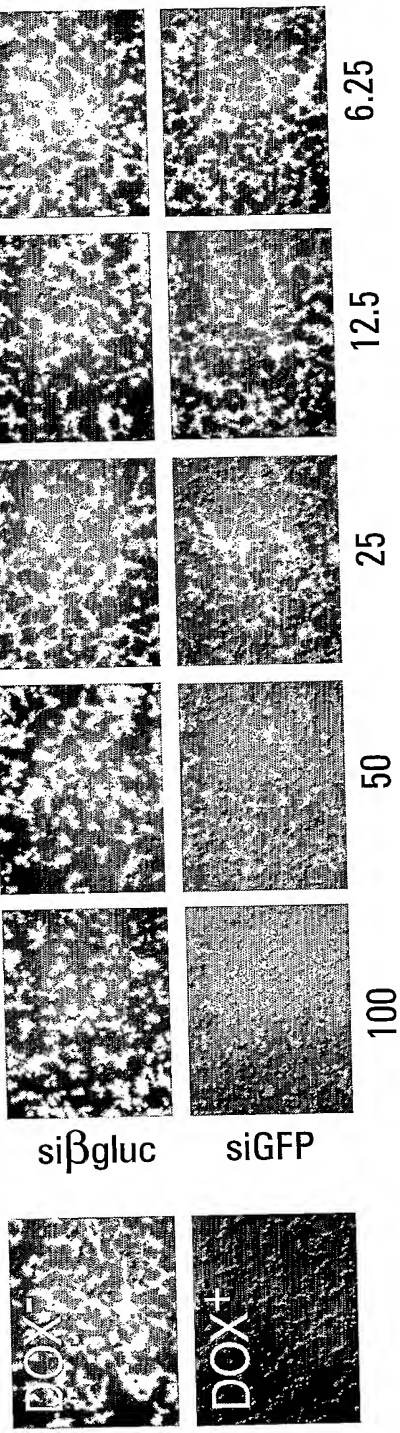


Fig. 3A

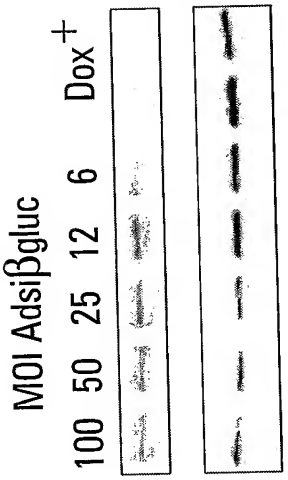


Fig. 3B

Fig. 3C

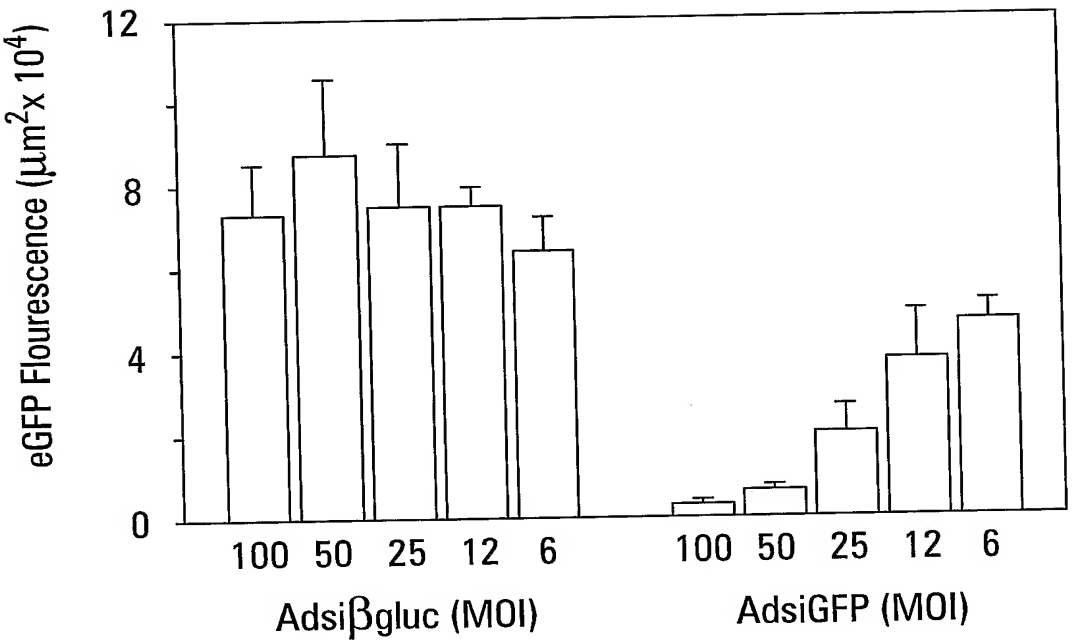

12
8
4
0
eGFP Flourescence (µm²x 10⁴)
100 50 25 12 6
Adsiβgluc (MOI)
100 50 25 12 6
AdsiGFP (MOI)

Fig. 3D

Fig. 4A

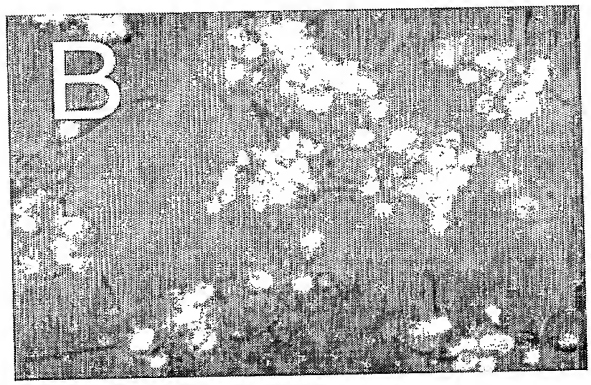


Fig. 4B

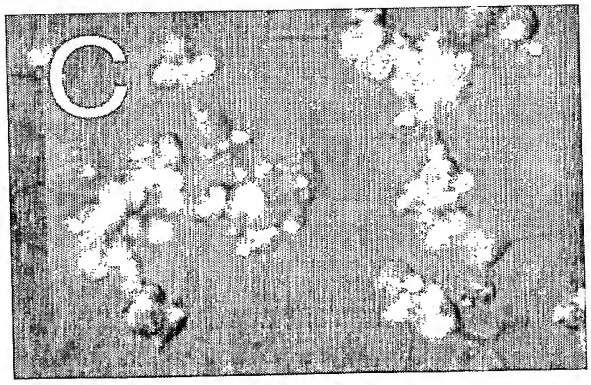


Fig. 4C

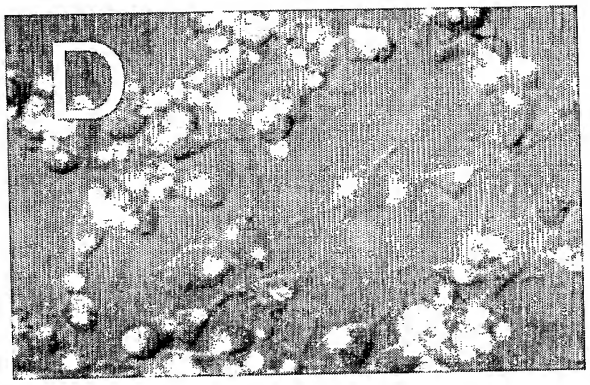


Fig. 4D

Fig. 4E

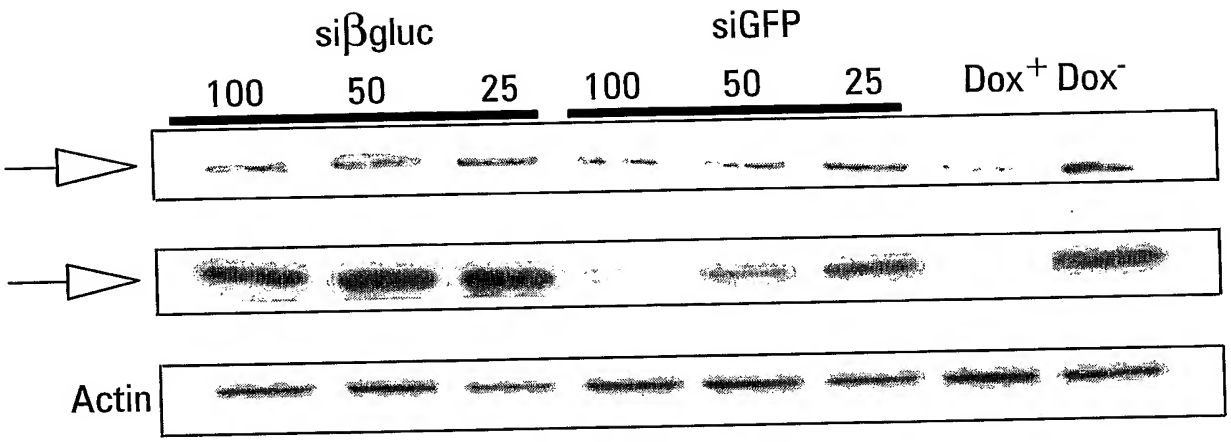

siβgluc
siGFP
100
50
25
100
50
25
Dox+ Dox-
Actin

Fig. 4F

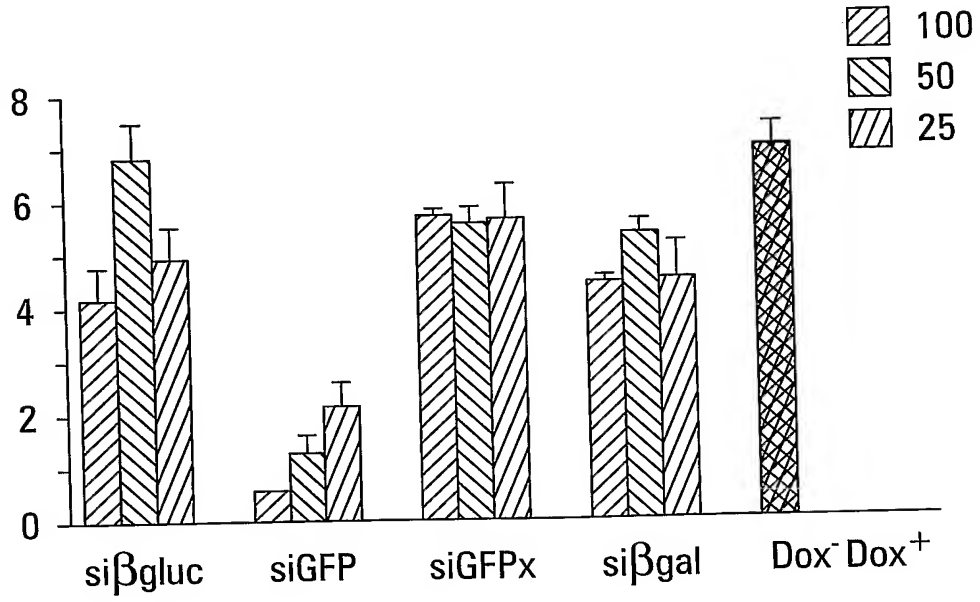

100
50
25
8
6
4
2
0
siβgluc
siGFP
siGFPx
siβgal
Dox- Dox+

Fig. 4G

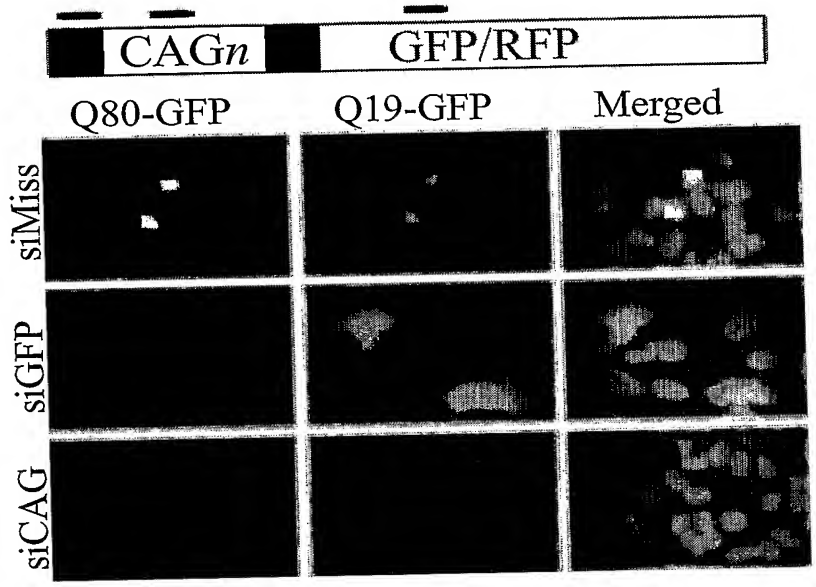

CAGn
GFP/RFP
Q80-GFP
Q19-GFP
Merged
siMiss
siGFP
siCAG

Fig. 5A

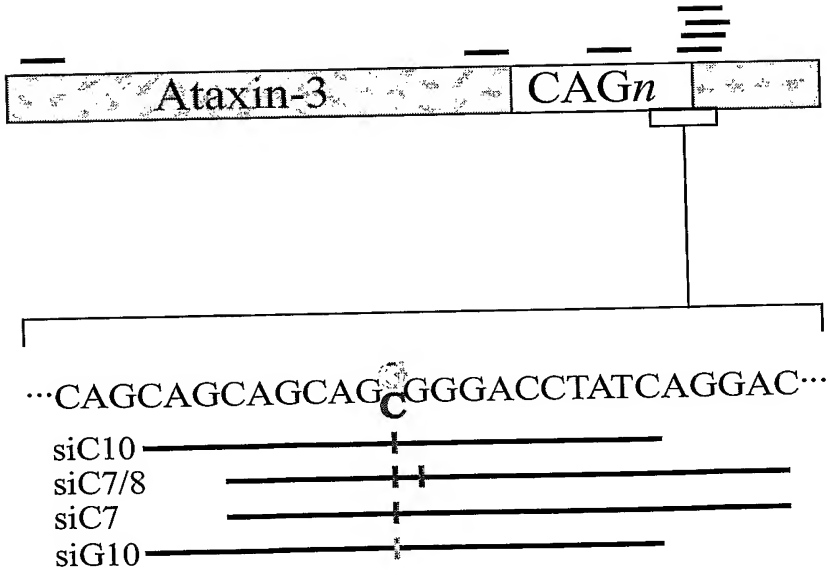

Ataxin-3
CAGn
···CAGCAGCAGCAGCGGGACCTATCAGGAC···
siC10
siC7/8
siC7
siG10

Fig. 5B

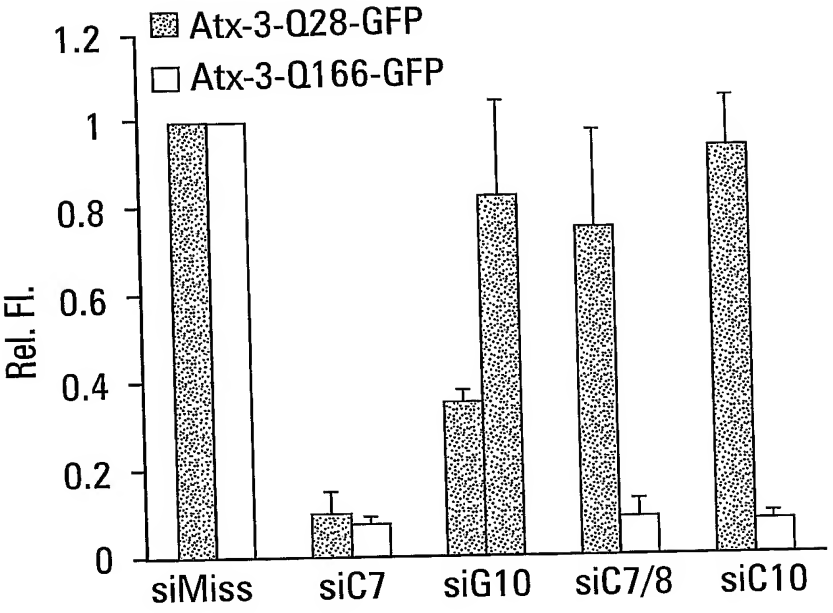

Atx-3-Q28-GFP
Atx-3-Q166-GFP
Rel. Fl.
1.2
1
0.8
0.6
0.4
0.2
0
siMiss
siC7
siG10
siC7/8
siC10

Fig. 5C

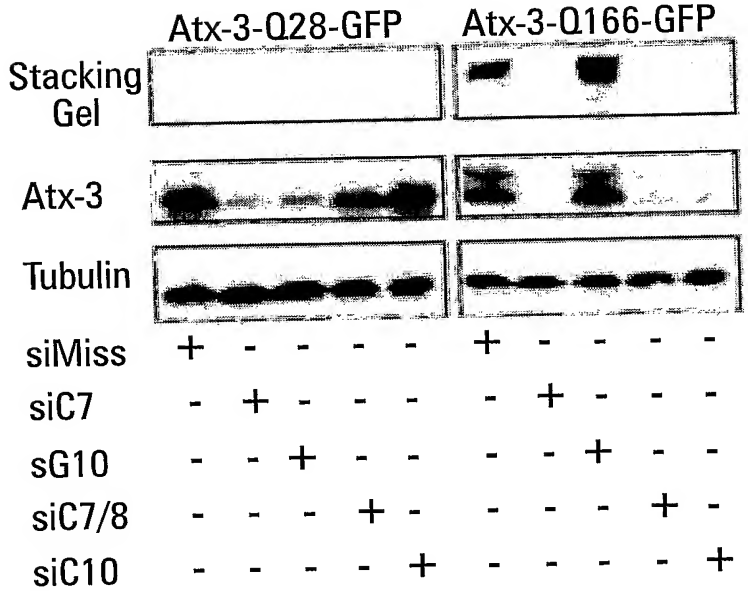

Atx-3-Q28-GFP
Atx-3-Q166-GFP
Stacking Gel
Atx-3
Tubulin
siMiss + - - - - + - - - -
siC7 - + - - - - + - - -
sG10 - - + - - - - + - -
siC7/8 - - - + - - - - + -
siC10 - - - - + - - - - +

Fig. 5D

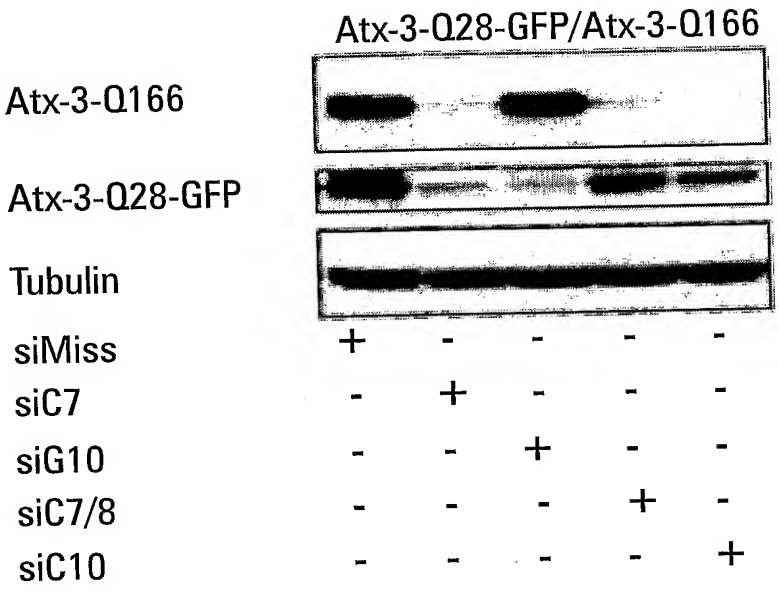

Atx-3-Q28-GFP/Atx-3-Q166
Atx-3-Q166
Atx-3-Q28-GFP
Tubulin
siMiss + - - - -
siC7 - + - - -
siG10 - - + - -
siC7/8 - - - + -
siC10 - - - - +

Fig. 5E

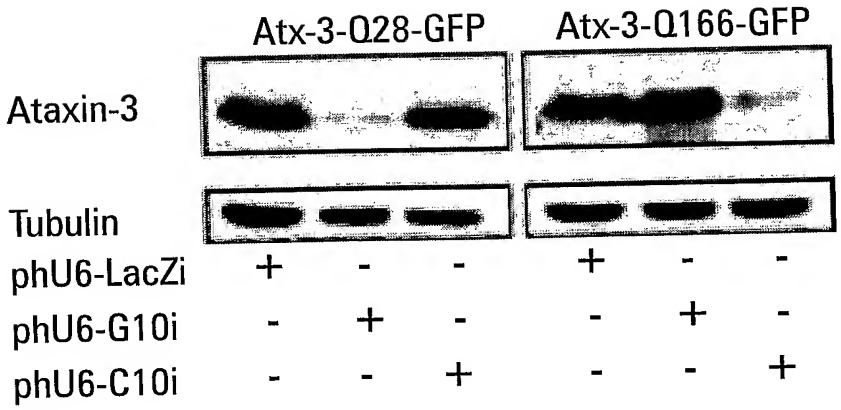

Atx-3-Q28-GFP
Atx-3-Q166-GFP
Ataxin-3
Tubulin
phU6-LacZi + - - + - -
phU6-G10i - + - - + -
phU6-C10i - - + - - +

Fig. 5F

Ataxin-3		Tau	
NAME	Primer Sequence (5'-3')	NAME	Primer Sequence (5'-3')
siMiss	CGGCAAGCTGCGCATGAAGTTC ATGAACTTCATGCTCAGCTTGC	siN'-Tau	TCGAAGTGATGGAAGATCACGC CAGCGTGATCTTCCATCACTTC
siGFP	ATGAACTTCAGGGTCAGCTTGC CGGCAAGCTGACCCTGAAGTTC	si272	CAGCCGGGAGTCGGGAAGGTGC CTGCACCTTCCCGACTCCCGGC
siC7	CAGCAGCGGGACCTATCAGGAC CTGTCCTGATAGGTCCCGCTGC	si301	ACGTCCTCGGCGGCGGCAGTGTGC TTGCACACTGCCGCCTCCGCGGAC
siG10	CAGCAGCAGGGGGACCTATC CTGATAGGTCCCCCTGCTGC	si406	ACGTCTCCATGGCATCTCAGC TTGCTGAGATGCCATGGAGAC
siC7/8	CAGCAGCCGGACCTATCAGGAC CTGTCCTGATAGGTCCGGCTGC	siA9	GTGGCCAGATGGAAGTAAAATC CAGATTTTACTTCCATCTGGCC
siC10	CAGCAGCAGCGGGACCTATC CTGATAGGTCCCGCTGCTGC	siA9/C8	GTGGCCACATGGAAGTAAAATC CAGATTTTACTTCCATGTGGCC
siN'CAG	TTGAAAAACAGCAGCAAAAGC CTGCTTTTGCTGCTGTTTTTC	siA9/C12	GTGGCCAGATGCAAGTAAAATC CAGATTTTACTTGCATCTGGCC
siCAG	CAGCAGCAGCAGCAGCAGCAGC CTGCTGCTGCTGCTGCTGCTGC		

Fig. 6

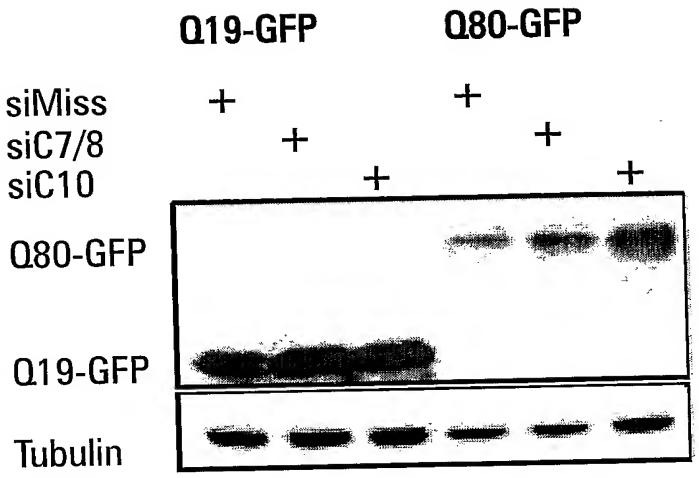

Q19-GFP
Q80-GFP
siMiss
siC7/8
siC10
Q80-GFP
Q19-GFP
Tubulin

Fig. 7A

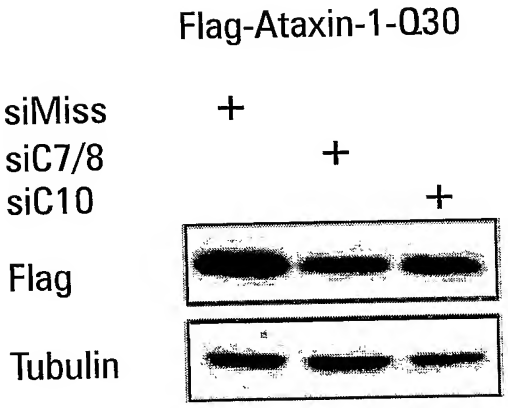

Flag-Ataxin-1-Q30
siMiss
siC7/8
siC10
Flag
Tubulin

Fig. 7B

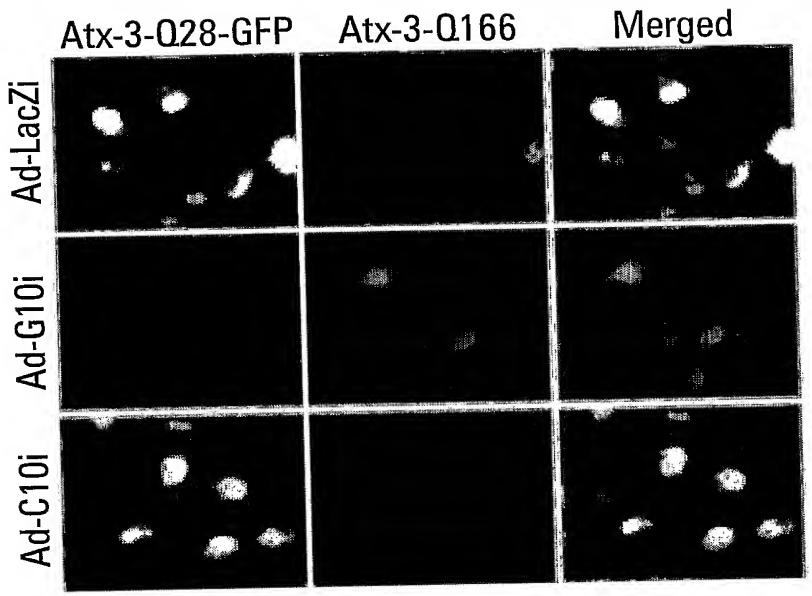

Atx-3-Q28-GFP
Atx-3-Q166
Merged
Ad-LacZi
Ad-G10i
Ad-C10i

Fig. 8A

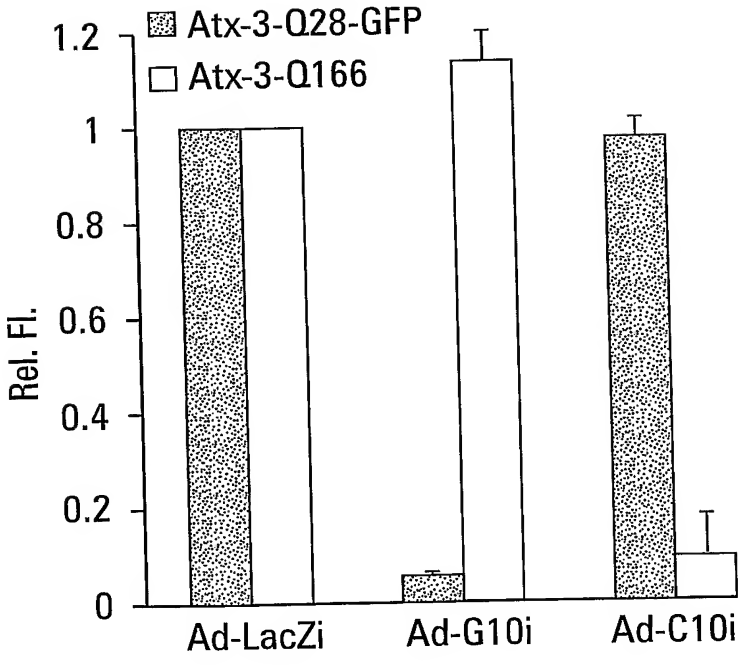

Atx-3-Q28-GFP
Atx-3-Q166
Rel. Fl.
1.2
1
0.8
0.6
0.4
0.2
0
Ad-LacZi
Ad-G10i
Ad-C10i

Fig. 8B

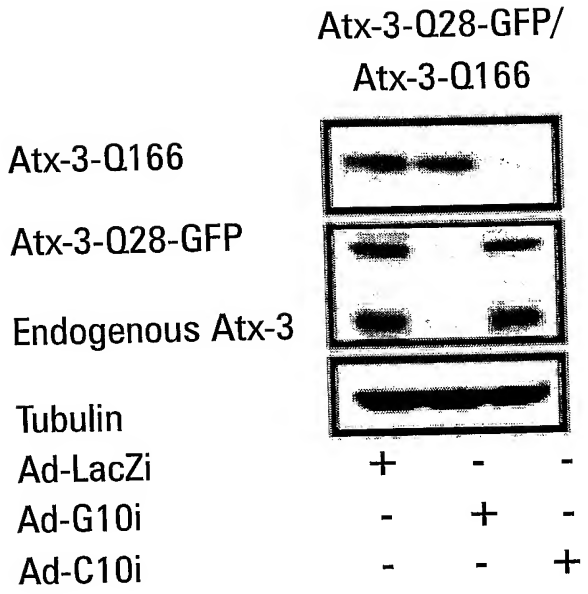

Atx-3-Q28-GFP/
Atx-3-Q166
Atx-3-Q166
Atx-3-Q28-GFP
Endogenous Atx-3
Tubulin
Ad-LacZi + - -
Ad-G10i - + -
Ad-C10i - - +

Fig. 8C

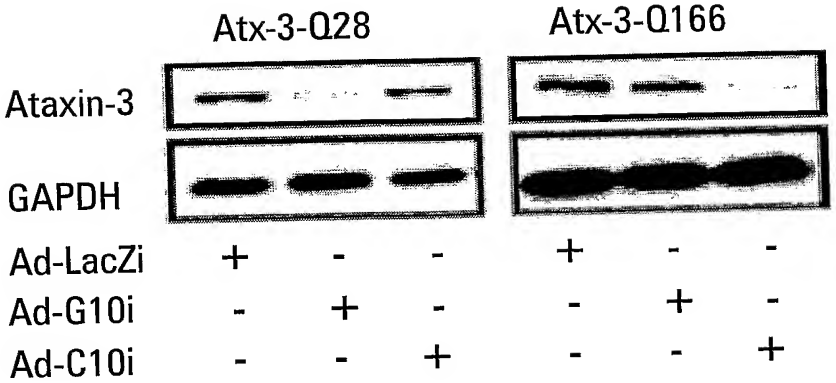

Atx-3-Q28
Atx-3-Q166
Ataxin-3
GAPDH
Ad-LacZi + - - + - -
Ad-G10i - + - - + -
Ad-C10i - - + - - +

Fig. 8D

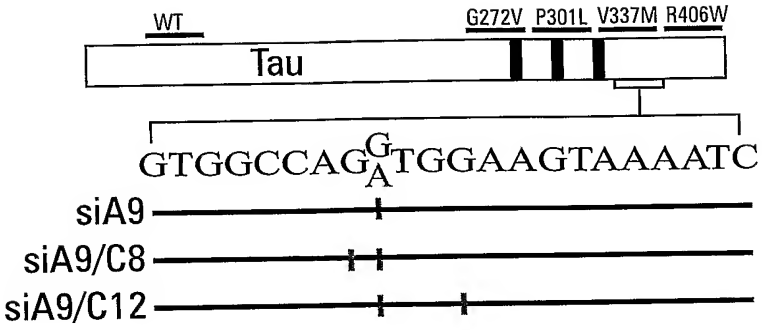


Fig. 9A

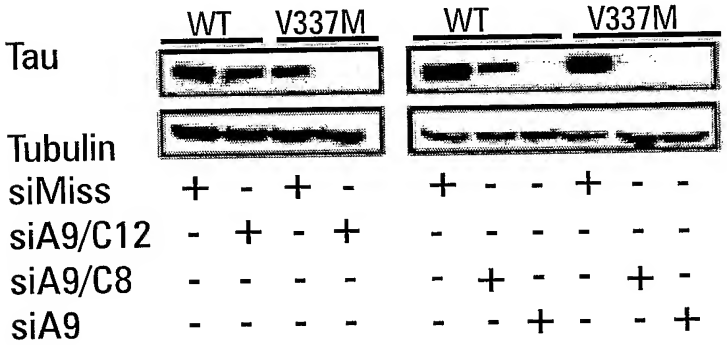


Fig. 9B

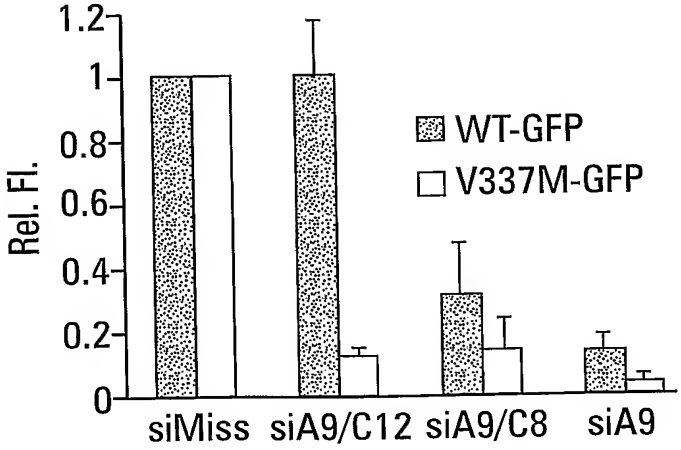


Fig. 9C

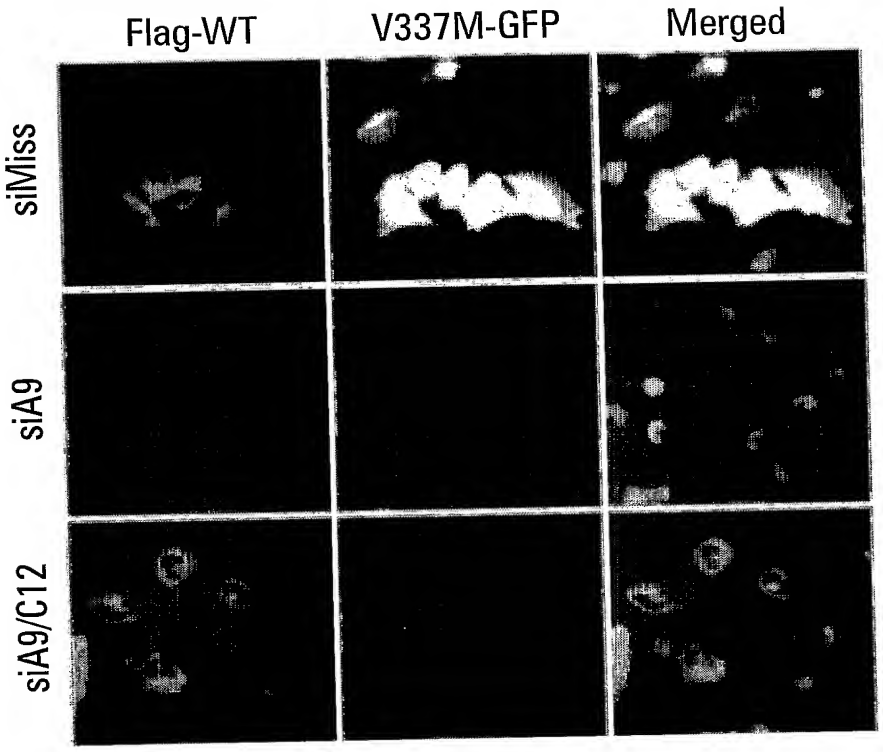


Fig. 10A

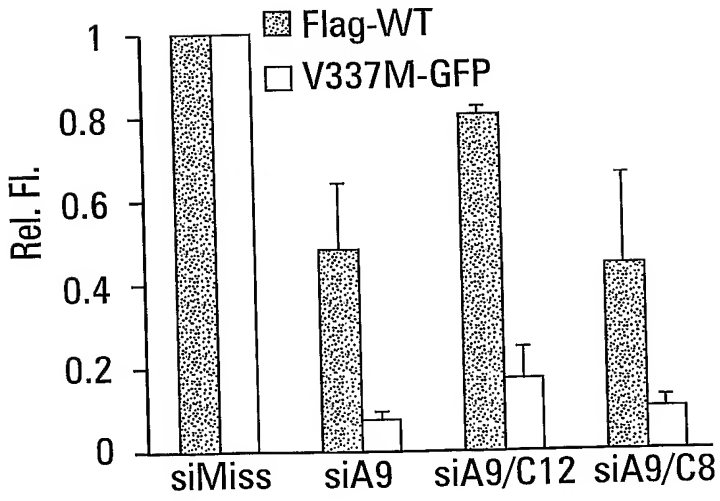


Fig. 10B

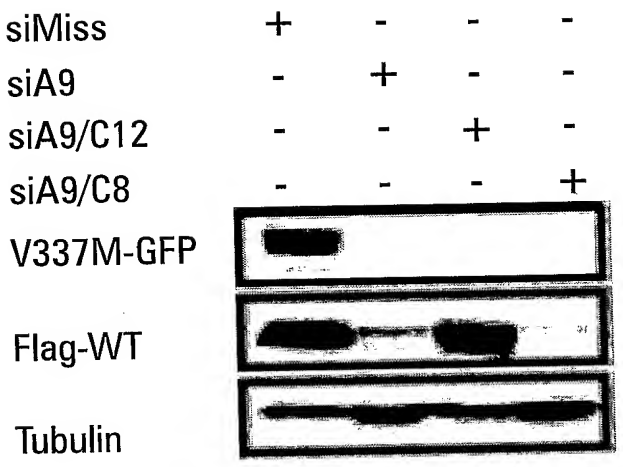


Fig. 10C

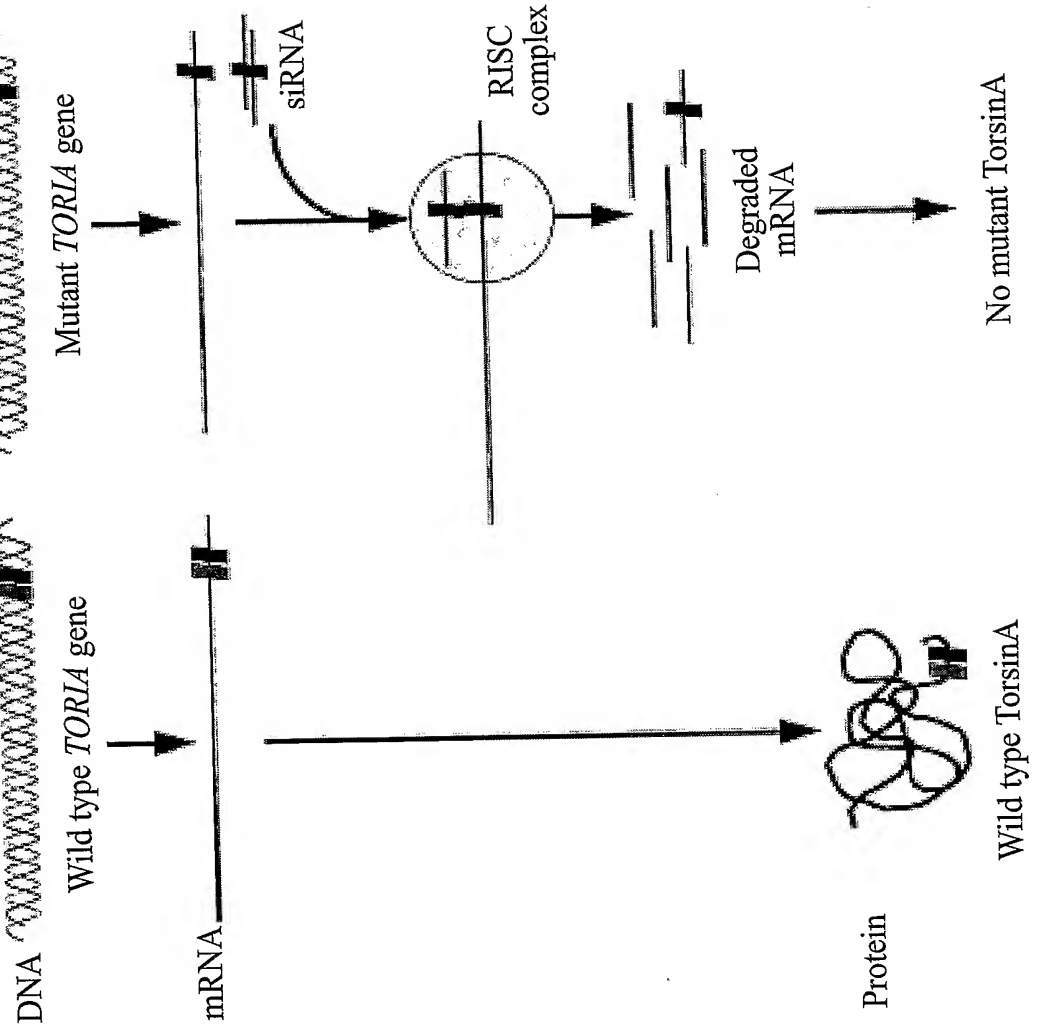

DNA
Wild type TORIA gene
mRNA
Protein
Wild type TorsinA
Mutant TORIA gene
siRNA
RISC complex
Degraded mRNA
No mutant TorsinA

Fig. 11

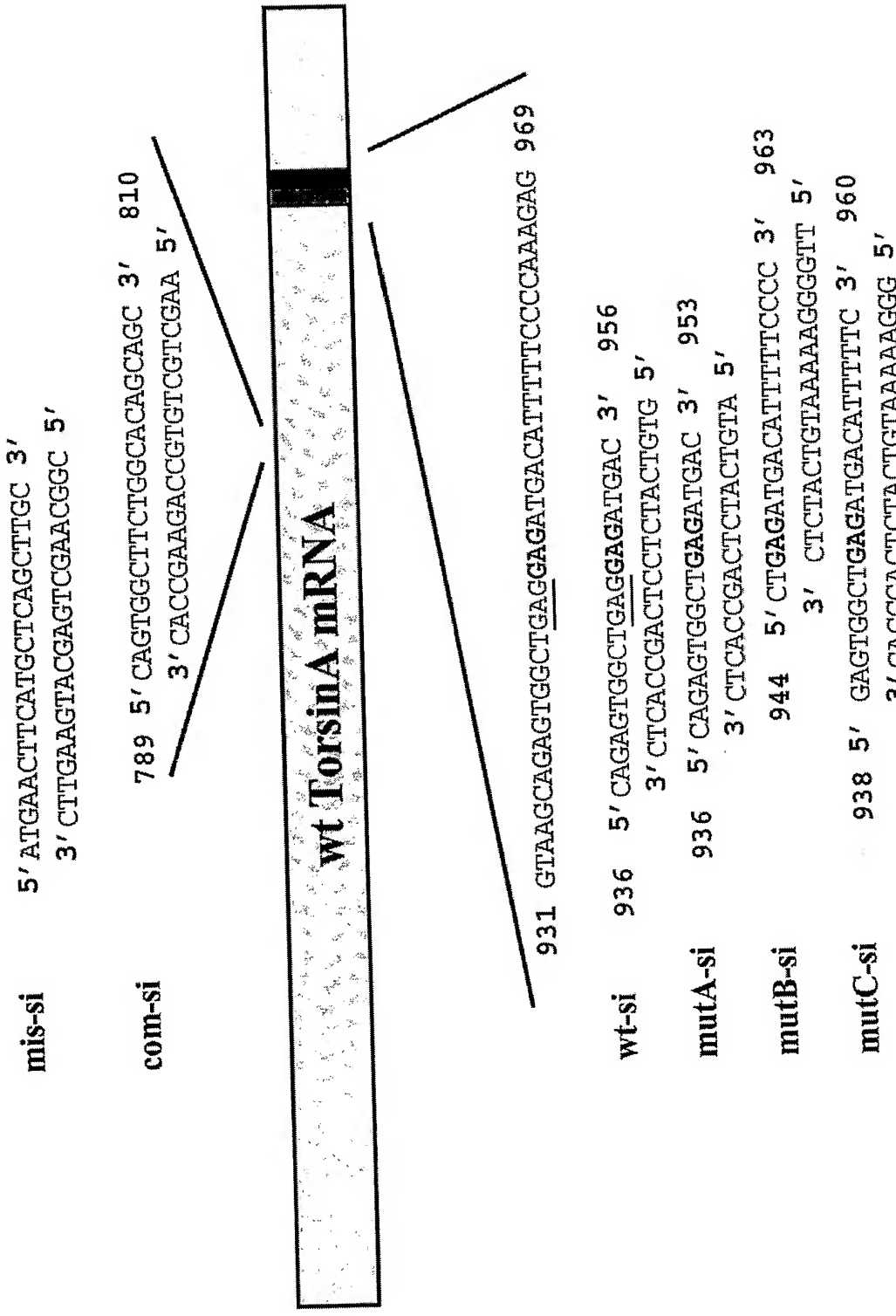

mis-si
5′ ATGAACTTCATGCTCAGCTTGC 3′
3′ CTTGAAGTACGAGTCGAACGGC 5′
com-si
789 5′ CAGTGGCTTCTGGCACAGCAGC 3′ 810
3′ CACCGAAGACCGTGTCGTCGAA 5′
wt TorsinA mRNA
931 GTAAGCAGAGTGGCTGAGGAGATGACATTTTTCCCCAAAGAG 969
wt-si 936 5′ CAGAGTGGCTGAGGAGATGAC 3′ 956
3′ CTCACCGACTCCTCTACTGTG 5′
mutA-si 936 5′ CAGAGTGGCTGAGATGAC 3′ 953
3′ CTCACCGACTCTACTGTA 5′
mutB-si 944 5′ CTGAGATGACATTTTTCCCC 3′ 963
3′ CTCTACTGTAAAAAGGGGTT 5′
mutC-si 938 5′ GAGTGGCTGAGATGACATTTTTC 3′ 960
3′ CACCGACTCTACTGTAAAAAGG 5′

Fig. 12

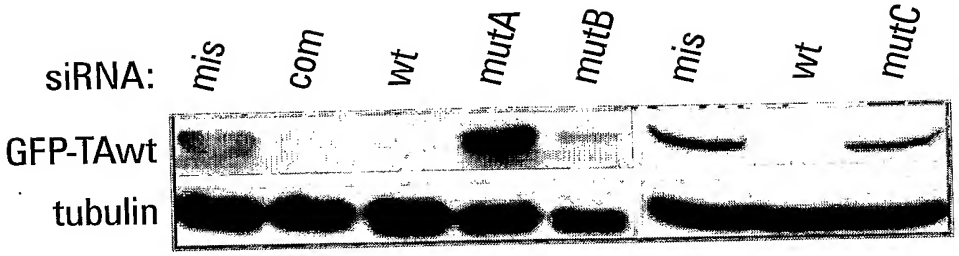


Fig. 13A

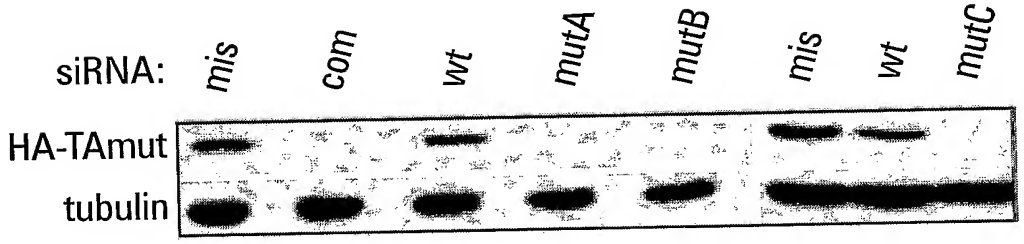


Fig. 13B

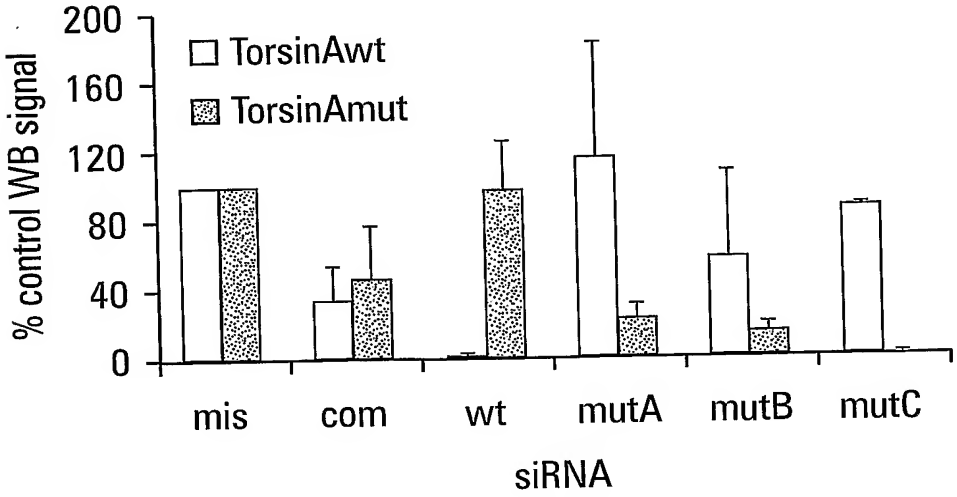


Fig. 13C

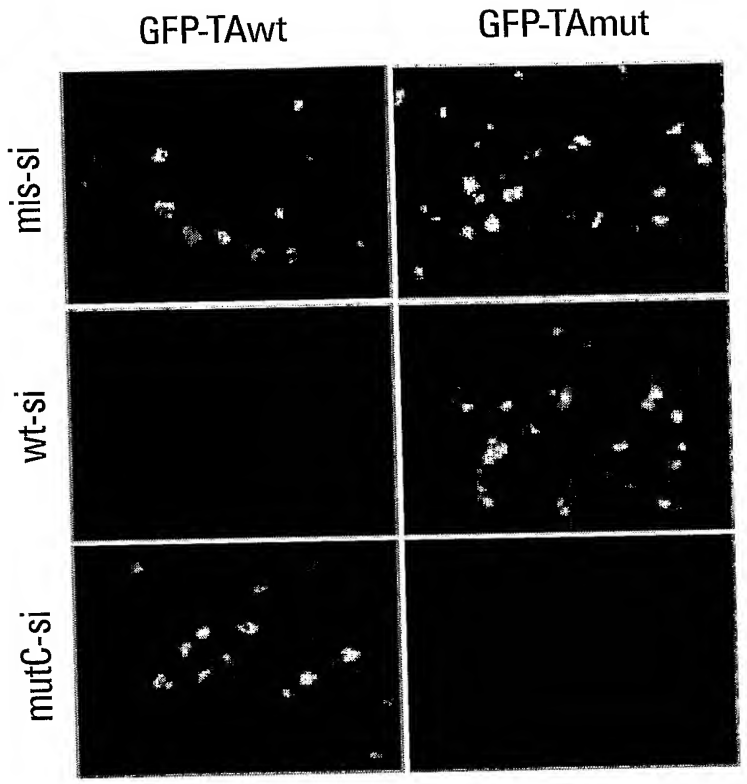


Fig. 13D

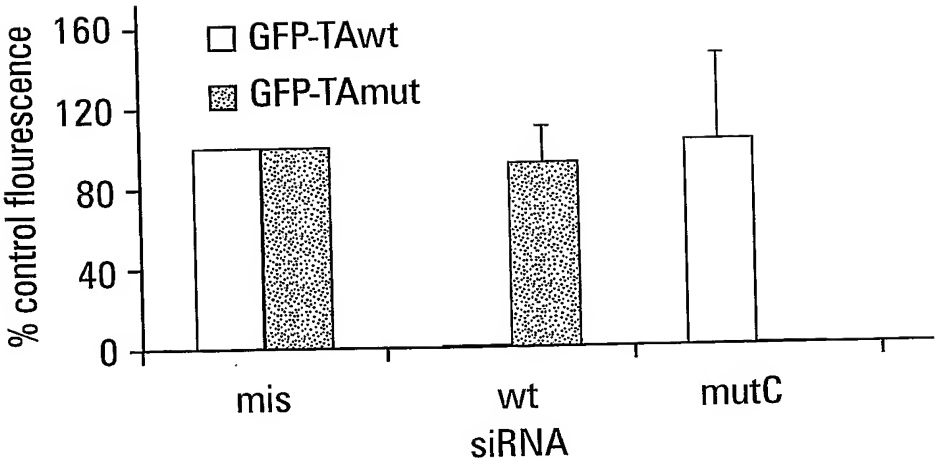


Fig. 13E

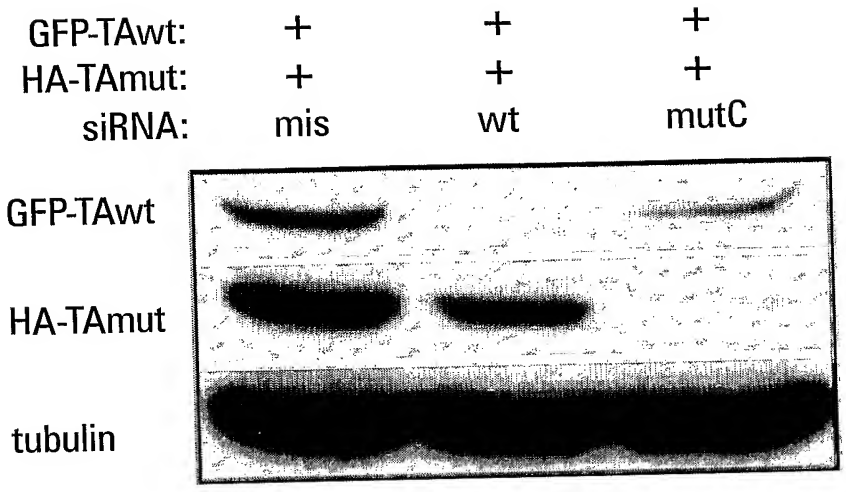


Fig. 14A

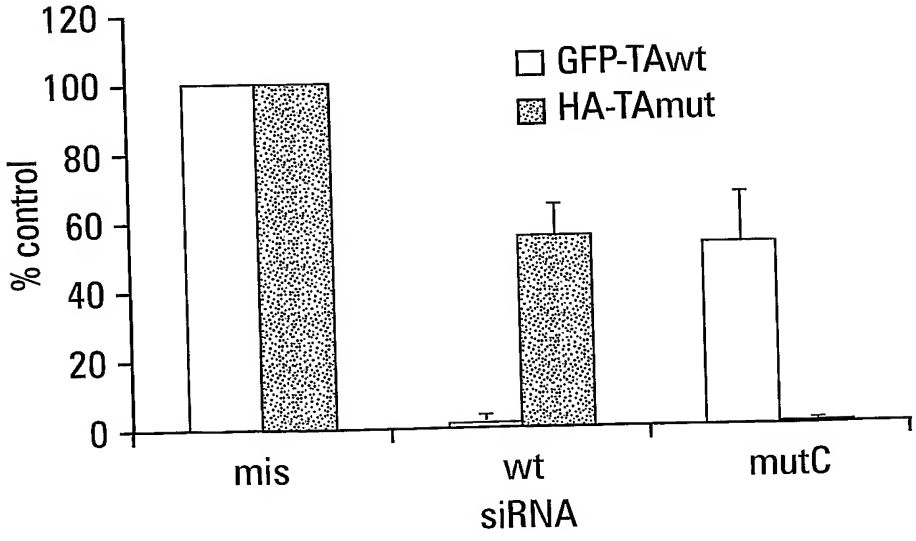


Fig. 14B

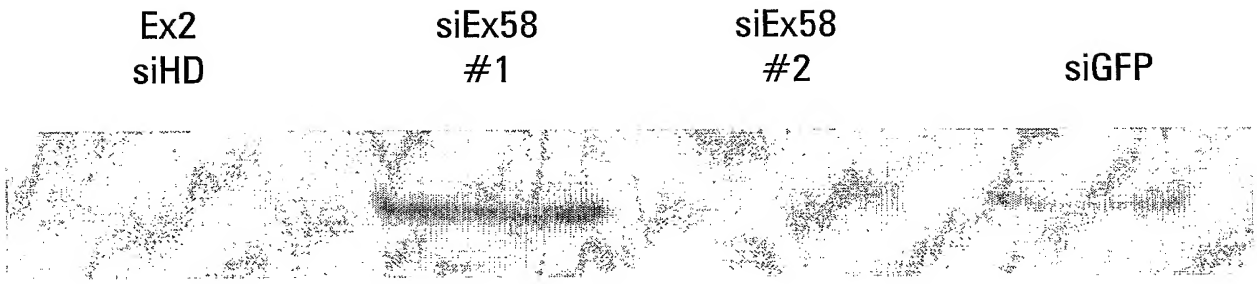


Fig. 15

Fig. 16A

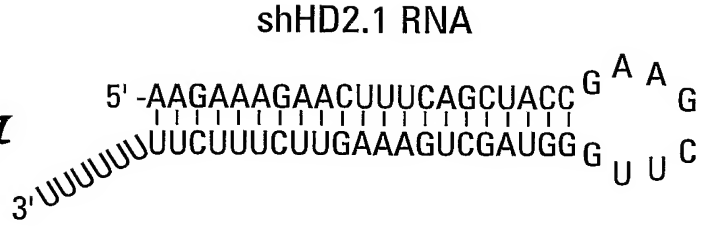


Fig. 16B

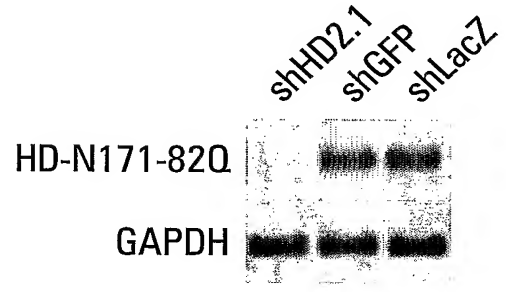


Fig. 16C

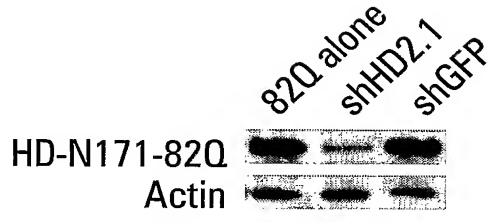


Fig. 16D

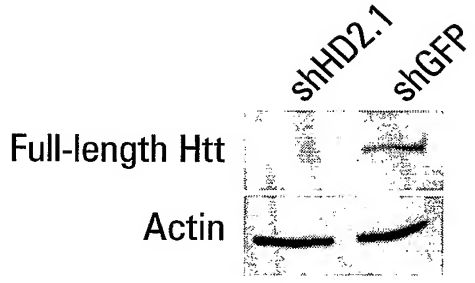


Fig. 16E

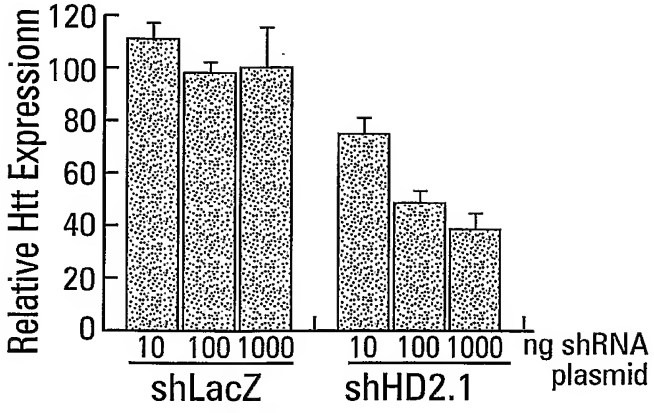


Fig. 17A

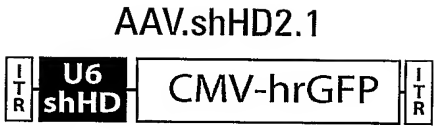


Fig. 17B

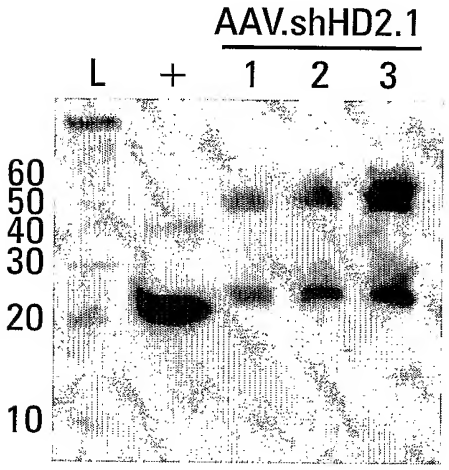


Fig. 17C

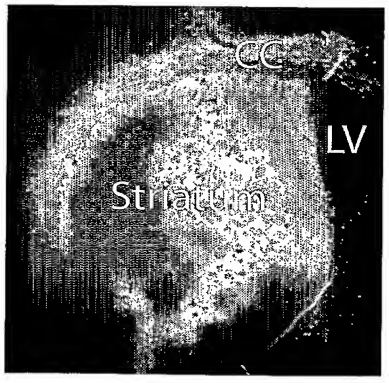
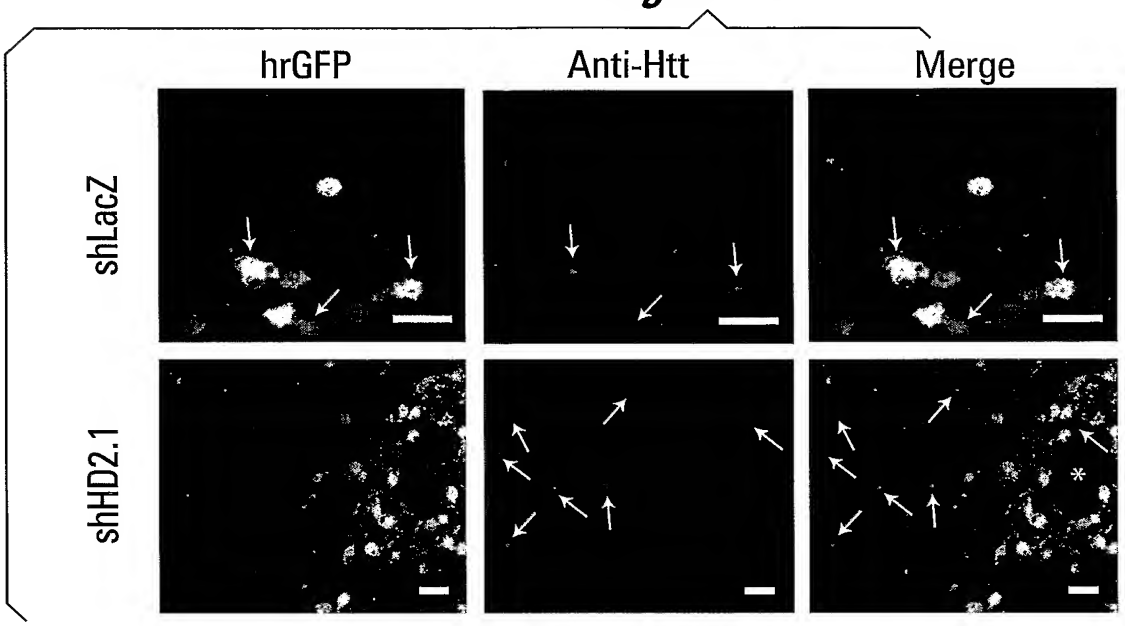


Fig. 18A



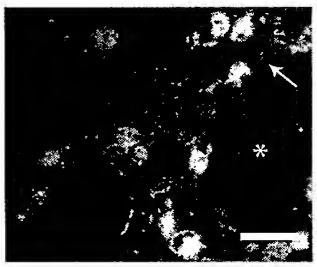


Fig. 18B

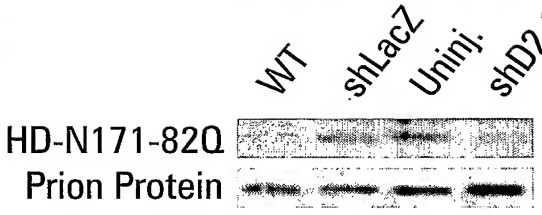


Fig. 18C

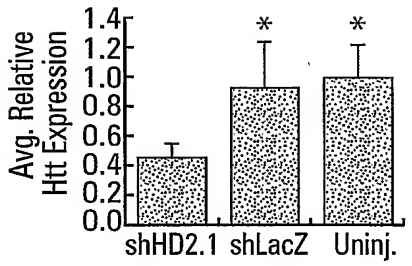


Fig. 18D

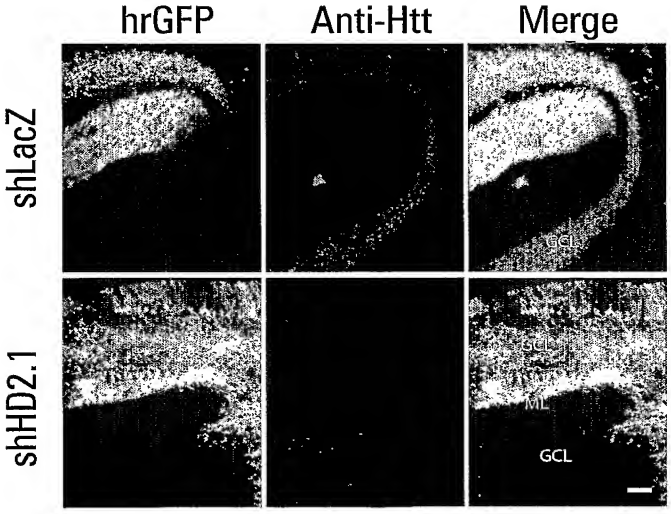


Fig. 18E

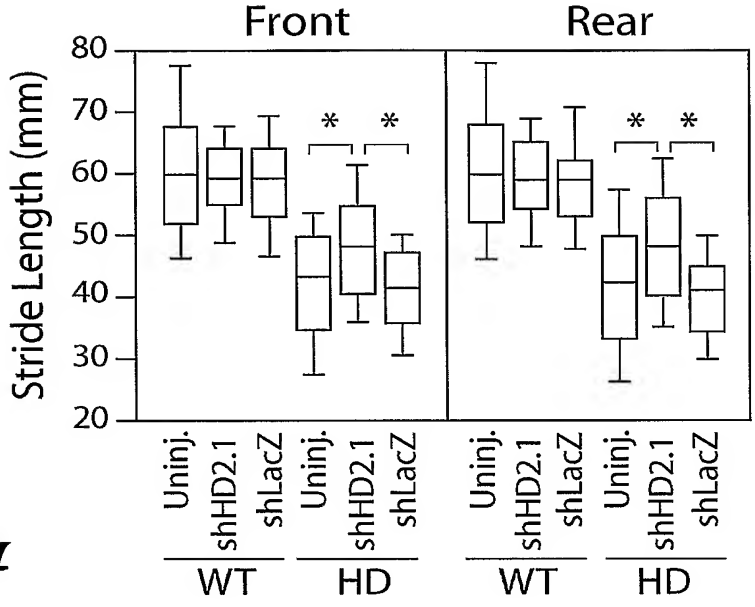


Fig. 19A

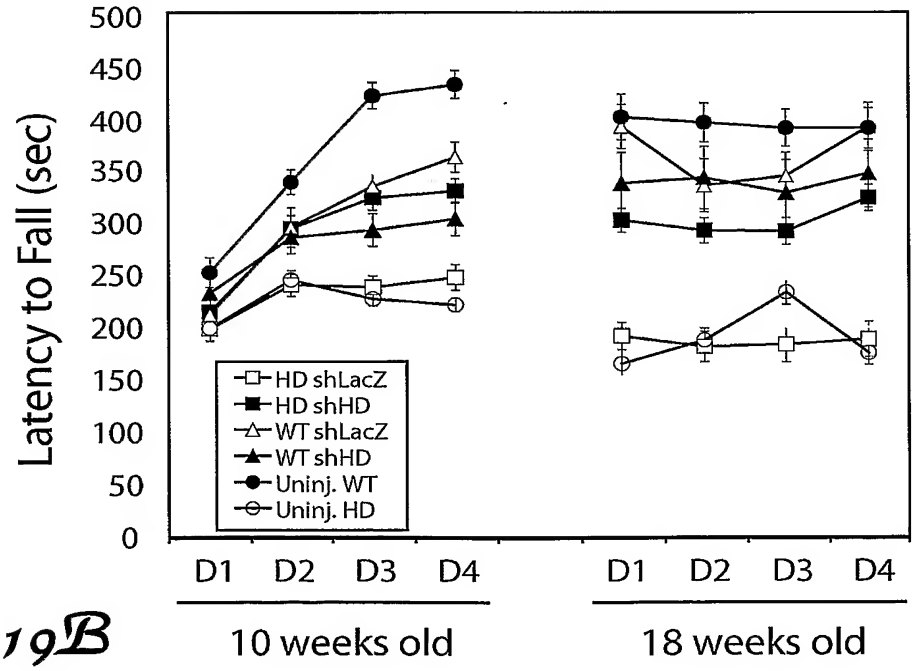


Fig. 19B

Fig. 20

Huntingtin hairpins

Harper PCR primer #	Read 5' to 3' Sense	Loop	Antisense
shHD2.1	AAGAAAGAACTTTCAGCTACC (SEQ ID NO:60)	GAAGCTTG (SEQ ID NO:61)	GGTAGCTGAAAGTTCTTTCTT (SEQ ID NO:62)
shHD2.2 (773)	AGAACTTTCAGCTACCAAG (SEQ ID NO:63)	CTTCCTGTCA (SEQ ID NO:64)	CTTGGTAGCTGAAAGTTCTTT (SEQ ID NO:65)
shHD3.1 (950)	TGCCTCAACAAAGTTATCA (SEQ ID NO:66)	CTTCCTGTCA	TGATAACTTTGTTGAGGCATT (SEQ ID NO:67)
348 Htt/Hdh #1	CAGCTTGTCCAGGTTTATGAA (SEQ ID NO:68)	CTTCCTGTCA	TTCATAAACCTGGACAAGCTG (SEQ ID NO:69)
350 Hdh#1	GACCGTGTGAATCATTGTCT**A** (SEQ ID NO:70)	CTTCCTGTCA	**T**AGACAATGATTCACACGGTC (SEQ ID NO:71)
351 Htt #1	TGGCACAGTCT**G**TCAGAAATT (SEQ ID NO:72)	CTTCCTGTCA	AATTTCTGA**C**AGACTGTGCCA (SEQ ID NO:73)
356 Htt/Hdh #2	CTGGAATGTTCCGGAGAATCA (SEQ ID NO:74)	CTTCCTGTCA	TGATTCTCCGGAACATTCCAG (SEQ ID NO:75)
352 Htt/Hdh #3	TTCTCTTCTGTGATTATGTCT (SEQ ID NO:76)	CTTCCTGTCA	AGACATAATCACAGAAGAGAA (SEQ ID NO:77)
354 Htt/Hdh #3	GTCCACCCCCTCCATCATTTA (SEQ ID NO:78)	CTTCCTGTCA	TAAATGATGGAGGGGGTGGAC (SEQ ID NO:79)
353 Htt/Hdh #4	AAGAAAGACCGTGTGAATCAT (SEQ ID NO:80)	CTTCCTGTCA	ATGATTCACACGGTCTTTCTT (SEQ ID NO:81)
355 Hdh #2	GGGCATCGCTATGGAACT**G**TT (SEQ ID NO:82)	CTTCCTGTCA	AA**C**AGTTCCATAGCGATGCCC (SEQ ID NO:83)
349 Htt #2	**G**CCGCTGCACCGACCAAAGAA (SEQ ID NO:84)	CTTCCTGTCA	TTCTTTGGTCGGTGCAGCGG**C** (SEQ ID NO:85)
358 Htt #3	**G**ACCCTGGAAAAGCTGATGAA (SEQ ID NO:86)	CTTCCTGTCA	TTCATCAGCTTTTCCAGGGT**C** (SEQ ID NO:87)
357 Htt/Hdh #5	AGCTTTGATGGATTCTAATCT (SEQ ID NO:88)	CTTCCTGTCA	AGATTAGAATCCATCAAAGCT (SEQ ID NO:89)

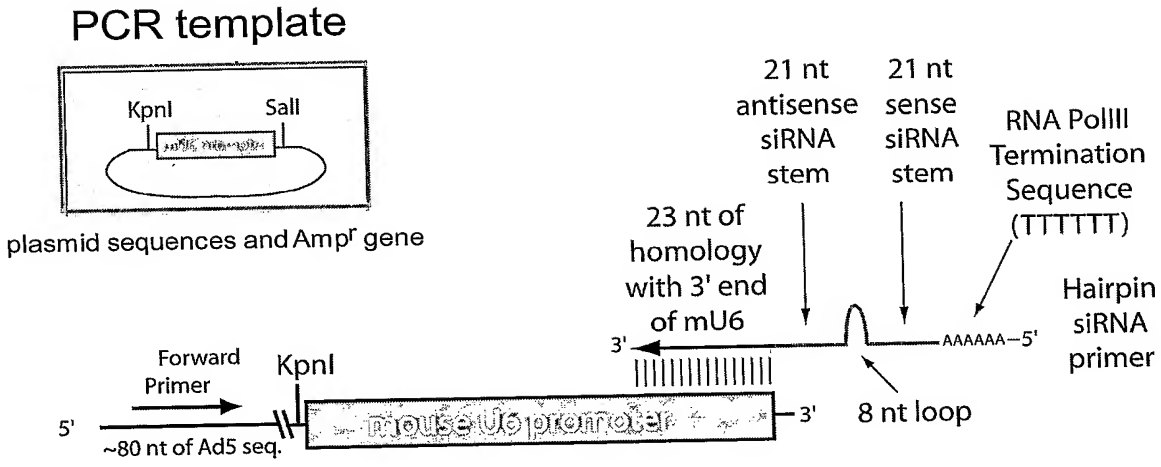

PCR template
KpnI
SalI
plasmid sequences and Ampr gene
21 nt antisense siRNA stem
21 nt sense siRNA stem
RNA PolIII Termination Sequence (TTTTTT)
23 nt of homology with 3' end of mU6
Hairpin siRNA primer
3'
AAAAAA—5'
Forward Primer
KpnI
5'
~80 nt of Ad5 seq.
mouse U6 promoter
3'
8 nt loop

Fig. 21

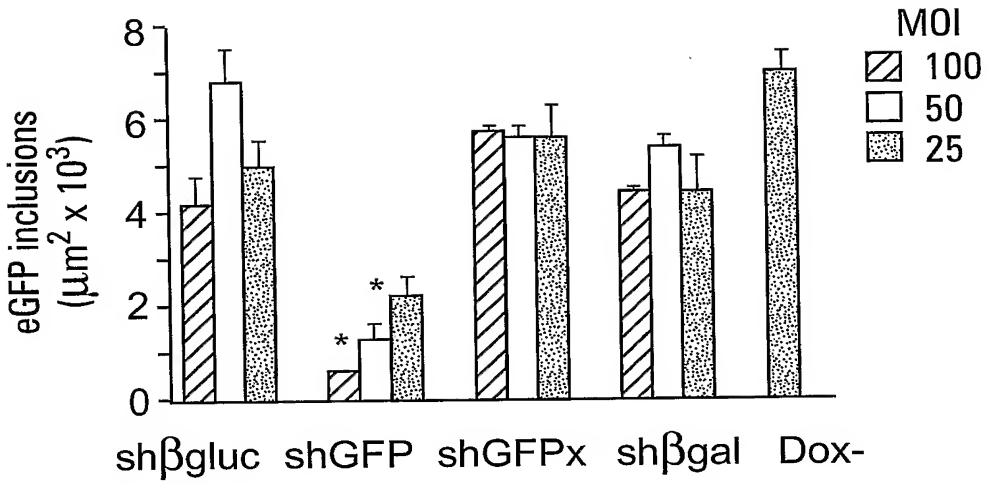

MOI
100
50
25
eGFP inclusions
(μm² x 10³)
8
6
4
2
0
*
*
shβgluc
shGFP
shGFPx
shβgal
Dox-

Fig. 22

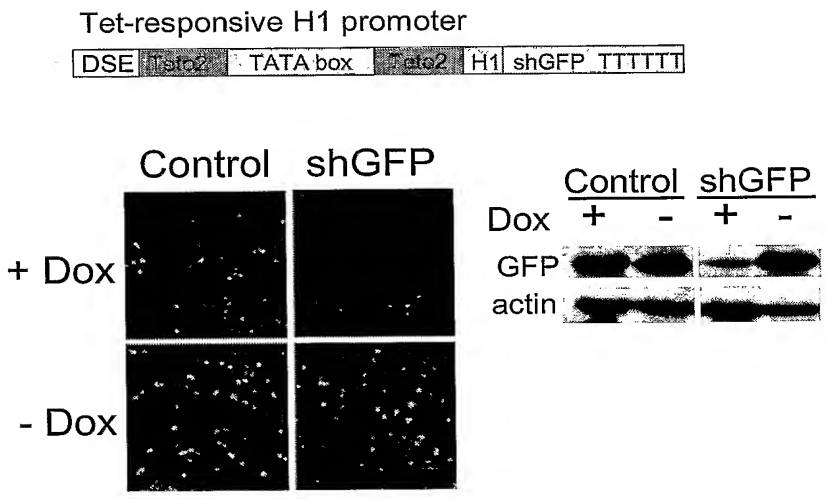

Tet-responsive H1 promoter
DSE
Tetο2
TATA box
Tetο2
H1
shGFP
TTTTTT
Control
shGFP
+ Dox
- Dox
Control
shGFP
Dox
+
-
+
-
GFP
actin

Fig. 23

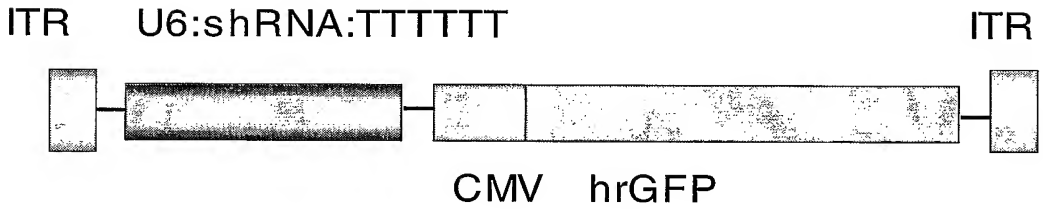


Fig. 24